Four Essays in Macroeconomic Modeling of Exchange Rates, Interest Rates and Commodity Prices

Dissertation
zur Erlangung der Würde eines
Doktors der Staatswissenschaften

vorgelegt der
Wirtschaftswissenschaftlichen Fakultät der
Universität Basel

von Matthias Gubler
von Wila, Zürich

Buchbinderei Beat Gschwind
Basel, 2012

Genehmigt von der Wirtschaftswissenschaftlichen Fakultät der Universität Basel auf Antrag von Prof. Dr. Peter Kugler und PD Dr. Sylvia Kaufmann.

Basel, den 13. April 2012 Der Dekan: Prof. Dr. George Sheldon

Acknowledgments

I would like to express my gratitude to all those who have given me guidance and support in the process of completing this project. First of all, I am indebted to my supervisor Peter Kugler for his support, encouragement and inspiring conversations. I would like to express my sincere gratitude and appreciation to him for giving me the unique opportunity to work as an assistant in his department. I am also much obliged to my second reader Sylvia Kaufmann who supported me, especially when I had to break through the obstacles involved in starting to work on a Ph.D. thesis. Furthermore, I would like to thank Hermione Miller-Moser, who provided excellent editorial assistance.

I am grateful that I had the opportunity to benefit from attending the one-year Ph.D. program "Swiss Program for Beginning Doctoral Students in Economics" at the Study Center Gerzensee, a foundation of the Swiss National Bank. The program not only allowed me to deepen my understanding of economics, but also gave me the opportunity to meet students from other Swiss universities. I would like to thank Michael Lobsiger, Marc Zahner and, in particular, Patrick Arni for many fruitful discussions on economics, and first and foremost for their friendship.

I am indebted to my co-authors Matthias S. Hertweck and Christoph Sax for providing an excellent research environment and for being a source of motivation. I am thankful for the numerous hours that we spend together to improve our research, and for our friendship.

Furthermore, I would like to thank Christian Glocker, Markus Ludwig, Iris Oberauner, Christian Rutzer and other friends from the faculty for their suggestions, help, interest, and valuable hints. I also want to thank my family, in particular my parents, and my friends for their support at all levels and – most of all – Linda, who gave me love and support.

Contents

Chapter 0

Introduction

This thesis consists of four chapters that examine different policy relevant economic processes related to exchange rates, interest rates and commodity prices. In particular, my thesis analyzes these topics empirically in the first three chapters, whereas the last chapter contains a model that can generate the empirical results of the third chapter. The second chapter is a study conducted jointly with Matthias S. Hertweck from the University of Konstanz (Germany). The third and fourth chapters are co-authored with Christoph Sax from the University of Basel.

The first chapter empirically investigates the relationship between speculators' currency carry trade positions and key financial variables which are of macroeconomic interest. The basic idea of a "currency carry trade" involves selling low-interest-rate currencies and investing simultaneously in high-interest-rate currencies. Investment strategies to exploit the failure of uncovered interest rate parity have become a major focus of interest. Therefore, carry trades also have appeared on policymakers' agendas. The analysis focuses on two target currencies, the US dollar and the euro, for which the Swiss franc serves as the funding currency. Since preliminary analyses point to regime-dependency with the interest-rate differential (IRD) as threshold variable, a multivariate threshold model is estimated and we account for conditional heteroscedasticity. Generalized impulse response functions differ in magnitude and significance between periods with a large and small IRD. Among others, dynamic responses indicate that in periods with a large IRD a positive shock to the IRD itself is not enough to compensate investors for the increased crash risk. In general, we find that carry trade positions are driven to a large extent by the expected risk in financial markets and the exchange rate. Furthermore, liquidity constraints can be important too, whereas the carry itself plays only a minor role. In addition, a sudden unwinding of carry trades has a significant impact on the nominal exchange rate, independent of the size of the interest-rate differential. Finally, Granger causality tests reveal that past position data help to predict exchange rate movements in periods with small interest-rate differentials, but feedback trading seems even more important.

The second chapter evaluates the importance of commodity price shocks in the U.S. business cycle. Thus, we extend the standard set of identified shocks in a 9-dimensional SVAR to include unexpected changes in commodity prices. The standard set of identified shocks covers the often considered neutral technology shocks, investment-specific technology shocks, and monetary policy shocks. We aim to quantify the relative importance of these shocks in the U.S. business cycle. The key result is that commodity price shocks are a very important driving force of the U.S. business cycle, second only to investment-specific technology shocks. In particular, commodity price shocks are the main determinant of cyclical movements in (headline) inflation. Moreover, the historical decomposition of shocks indicates that commodity price shocks contribute most to the high degree of macroeconomic volatility in the 1970s, particularly during and after the first OPEC oil crisis, and are also an important determinant of the double-dip in the early 1980s, the economic boom in the early 1990s, and the short early 2000s recession. The results of a counterfactual exercise suggest that the Fed's systematic contractionary response achieves price stability in the long run, yet at the cost of a significant economic downturn in output and per-capita hours. Furthermore, business cycle fluctuations in output and per-capita hours are primarily driven by unexpected changes in the relative price of investment goods. Neutral technology shocks and monetary policy shocks, on the other hand, seem less relevant in explaining business cycle movements in key macroeconomic variables. At low frequencies, however, neutral technology shocks do play an important role in explaining output variability. Finally, we show that the estimation bias caused by low-frequency movements in the data becomes less important in a model with sufficient information.

The third chapter examines the robustness of the Balassa-Samuelson (BS) hypothesis, a widespread explanation for structural deviations from purchasing power parity (PPP). According to the hypothesis, price level differences between countries, expressed in the same currency, can be ascribed to different productivity differentials between the non-tradable and tradable sector. We apply a panel cointegration model to estimate the long-run relationship between the real exchange rate and key explanatory variables of OECD countries. The data set includes new sectoral total factor productivity (TFP) values constructed by the OECD. We find a negative relationship between the productivity in the tradable sector and the long-run real exchange rate for the last two decades. This result not only contradicts the BS hypothesis but also challenges the findings of previous research in favor of the BS hypothesis. This negative relationship is robust against the choice of the country sample, the start of the sample period, the model specification, the inclusion of additional explanatory variables and non-tradable productivity. Furthermore, the finding is confirmed when the TFP values are substituted by labor productivity (LP) values. On the other hand, the connection between the productivity of non-tradables and the real exchange rate seems not robust.

The fourth chapter sketches a static general-equilibrium framework that shows how skill-biased technological change may reverse the classic BS effect, leading to a negative relationship between the productivity in the tradable sector and the real exchange rate. There are two sectors in a small open economy, each producing a homogeneous good, tradable export goods and non-tradable services. In the tradable sector, low-skilled labor, together with capital, is used to produce an intermediate routine task good, which in turn is combined with high-skilled labor to produce the final tradable good. A key feature of the model is the substitutability of low-skilled labor and capital. In order to analyze the reversion of the BS effect, we assess the effect of an increase in capital productivity on the real exchange rate. An increase in capital productivity affects the real exchange rate through the demand for low-skilled labor in two ways: (1) a *labor-repellent* effect in the tradable sector if the elasticity of substitution between low-skilled labor and capital is high relative to the importance of the intermediate routine task good; and (2) a *labor-attracting* effect in the non-tradable sector. The opposite BS effect occurs if the labor-repellent effect in the tradable sector outweighs the labor-attracting effect in the non-tradable sector. The labor-repellent effect dominates the labor-attracting effect if the labor force of the tradable sector is large relative to the labor force of the non-tradable sector. The lower overall low-skilled labor demand diminishes the wage rate and thus the price level. Therefore, an increase in the productivity of tradables may be connected with a real exchange rate depreciation.

Chapter 1

Carry Trade Activities: A Multivariate Threshold Model Analysis

Abstract

In this empirical study we analyze the relationship between carry trade positions and some key financial as well as macroeconomic variables using a multivariate threshold model. It is often stated that the Swiss franc serves as a funding currency. Therefore, we focus on carry trades based on the currency pairs US dollar/Swiss franc and euro/Swiss franc. Generalized impulse responses differ in magnitude and significance between periods with a large and small interest-rate differential. Furthermore, in periods with a small interest-rate differential, carry trade activities "Granger-cause" the nominal exchange rate. The Granger causality test results further indicate feedback trading. Overall, carry trade positions are driven to a large extent by the expected risk in financial markets and the nominal exchange rate. Liquidity constraints can also be important, whereas the carry itself plays only a minor role.

Keywords: Carry Trades, Multivariate Threshold Model, Tsay Test, Generalized Impulse Response Functions, Bootstrap Method, Granger Causality

1 Introduction

In this chapter, we empirically investigate the relationship between speculators' currency carry trade positions and key financial variables which are of macroeconomic interest. The basic idea of a "currency carry trade" (hereinafter "carry trade") involves selling low-interest-rate currencies (e.g., by borrowing money) and investing simultaneously in high-interest-rate currencies. Low-interest-rate currencies, such as the Swiss franc or the Japanese yen, are called funding currencies, whereas high-yielding currencies are called target currencies.

Recently, investment strategies to exploit the failure of uncovered interest rate parity (UIP) have become a major focus of interest not only for financial market participants; carry trades also have appeared on policymakers' agendas, specifically on those of central bankers. For instance, Jean-Pierre Roth, former president of the governing board of the Swiss National Bank, pointed out the crucial role of carry trades in determining the nominal exchange rate in the medium run (Roth, 2007). In our analysis, we focus on two target currencies for which the Swiss franc (CHF) serves as the funding currency: the US dollar (USD) and the euro (EUR).

UIP states that the gains due to interest-rate differentials (IRDs) are offset by the loss arising in the depreciation of the target currency. However, several empirical studies emphasize the violation of UIP ("forward premium puzzle").[1] Meese and Rogoff (1983) compare the out-of-sample forecast accuracy of different structural exchange rate models and conclude that exchange rates follow a "near random walk". In fact, Fama (1984) shows that on average the target currency appreciates. This empirical anomaly of the foreign exchange market makes carry trades profitable on average.

While an extensive body of the literature on carry trades examines their profitability, the main contribution of this study is the empirical investigation of the interaction between carry trade activities and financial as well as macroeconomic variables with a multivariate threshold model. Carry traders presumably react to shocks to variables which determine the profitability of their investment strategy, such as the interest-rate differential (the so-called "carry"), the nominal exchange rate, the risk sentiment, the investment return, and possible liquidity constraints. In addition, these variables can move due to unexpected carry trade activities. Thus, we include these variables, or reasonable proxies, in our model.

Therefore, our empirical study is closest to Brunnermeier et al. (2009) and Nishigaki (2007). Brunnermeier et al. (2009) show that in times of reduced funding liquidity and declining risk appetite carry traders are subject to crash risk due to the sudden unwinding of carry trades. Nishigaki (2007) examines the yen carry trade. His analysis implies that the carry has no significant impact on carry trade movements, in contrast to US stock

[1]For a literature survey, see for example Engel (1996).

prices. The results also indicate USD depreciation against the Japanese yen once carry trades unwind. Both of these studies incorporate futures positions to proxy carry trade activities, as we do for the CHF/USD exchange rate. Yet, futures position data with respect to the CHF/EUR exchange rate are not available. Hence, we employ for the Euro market the carry-to-risk ratio (CTR ratio) to proxy carry trade activities, since it is an important indicator of potential carry trade profitability.

Recent studies highlight the importance of regime-dependent results (see Section 2), and indeed, preliminary analyses of the IRD indicate a nonlinear relationship among the variables in our model. The results of a Tsay (1998) test confirm the assumption of nonlinearity. Therefore, we apply a multivariate threshold model to account for the possible changes in the dynamic behavior of carry trade activities dependent on the size of the IRD.

By analyzing the generalized impulse response functions (GIRFs), we find the following main results: First, carry trade positions are driven to a large extent by the expected risk in financial markets and the exchange rate. Since the responses of all other variables to shocks depend on the size of the carry, these differences are carried over to the speculators' carry trade positions. The results indicate that in times with a large carry a positive one-standard deviation shock to the carry itself is not enough to compensate investors for the increased crash risk. Moreover, in line with the prediction of UIP, the CHF appreciates instantaneously against the USD in times with high IRDs, but not in the regime with low IRDs. Second, liquidity constraints can be important too, whereas the carry itself plays only a minor role. Third, a sudden unwinding of carry trades has a significant impact on the nominal exchange rate, independent of the size of the IRD. Finally, we show that the majority of impulse responses is similar for the CHF/USD and CHF/EUR exchange rates, although the proxy for carry trade positions differs.

Klitgaard and Weir (2004) analyze futures position data and state that net positions do not seem to "Granger-cause" the exchange rate movements of the following week.[2] We follow their approach and apply the Granger causality test to our regime-dependent model and find that past position data help to predict exchange rate movements in periods with small IRDs. Additionally, in samples with the USD as target currency, the exchange rate has very high predictive power for carry trade activities, pointing to feedback trading.[3]

The remainder of this chapter is organized as follows. Section 2 contains an overview of the related literature. Data sources and variable definitions are presented in Section 3. In Section 4, we outline the methodology used for our empirical study. We provide a detailed discussion on our results for the GIRFs in Section 5 and their robustness analysis (Section 5.3). Section 5.4 shows the Granger causality test results and Section 6 concludes.

[2]See also Mogford and Pain (2006) for a similar study.

[3]In contrast, no prediction power is found in samples with the EUR as target currency. This might be due to the definition of the CTR ratio. This issue is discussed in greater detail in Section 5.4.

2 Related Literature

A large body of the literature on carry trades examines the profitability of potential carry trade strategies. A few studies conclude that these investment strategies lead to excess returns. These excess returns can be attributed neither to standard risk factors (Burnside et al., 2006), to the exposure to currency crashes (Jurek, 2007), nor to disaster risks (Farhi et al., 2009). Instead, market frictions such as the bid-ask spread and price pressure greatly reduce the return on these portfolios (Burnside et al., 2006), or they are not economically significant (Wagner, 2008). In contrast, Lustig et al. (2011) argue that carry trade profits are a compensation for systematic risk. Moreover, Darvas (2009) shows that the degree of leverage is crucial for excess returns. Profitability declines with increasing leverage. Furthermore, Kohler (2007) examines the correlation dynamics between returns on global equity portfolios and simple carry trade investment strategies. Based on his results, carry trades are exposed to a severe diversification meltdown in times of global stock markets crisis.

Another stream of the carry trade literature examines other channels to detect carry trade positions that focus mainly on yen carry trades. For example, Gagnon and Chaboud (2007) emphasize the "canonical yen carry trade" in contrast to the "derivatives carry trade" studied by Nishigaki (2007) and Brunnermeier et al. (2009).[4] Galati et al. (2007) compare low frequency data from the BIS international banking statistics with higher frequency futures data and find similar insights for carry trade positions. Cai et al. (2001) examine the effects of order flows and macroeconomic news on the dramatic yen/dollar volatility of 1998 with weekly data from the US Treasury on purchases and sales of spot, forward, and futures contracts. Finally, Hattori and Shin (2007) conclude that the waxing and waning of the balance sheets of foreign banks in Japan is related to the state of overall risk appetite. By using descriptive statistics and a simple econometric analysis, they reveal a positive relationship between the IRD[5] and carry trades. However, McGuire and Upper (2007) argue that carry trade positions are not only difficult to detect but also to distinguish from other investment strategies.

The importance of regime-dependent results is highlighted by Clarida et al. (2009) among others. These authors examine carry trade strategies and identify a robust empirical relationship between their excess returns and exchange rate volatility. Furthermore, they show that the failure of UIP is only present in low-volatility environments. Jordà and Taylor (2009) argue that more sophisticated conditional carry trade strategies exhibit more favorable payoffs. They adopt a nonlinear regime-dependent model approach and

[4] Gagnon and Chaboud (2007) define canonical carry trades as borrowing low-yielding currencies and investing the proceeds in high-interest-rate currencies. In contrast, derivatives carry trades are defined as taking on leveraged positions in derivatives markets. More on this issue is provided in Section 3.1.

[5] The IRD is the difference between the Japanese overnight rate and the average of the US, Euro-zone and Australia policy rates.

add the fundamental equilibrium exchange rate (FEER) to their model. In distinction to our study, they choose the threshold value exogenously. Christiansen et al. (2011) provide a factor model with regression coefficients dependent on market volatility and liquidity to assess carry trade strategies. In volatile periods the excess returns have much higher exposure to the stock market and also more mean reversion.

To the best of our knowledge, there is only one theoretical contribution in the literature that focuses specifically on carry trades. Plantin and Shin (2010) incorporate funding externalities and carry costs into their model to predict the classic price pattern "going up the stairs, and coming down in the elevator". The increase in carry trade positions is followed by abrupt stochastic reversals.

3 Data

3.1 Variables

We collected data to examine the Swiss franc (CHF) carry trade with the US dollar (USD) or the euro (EUR) as respective target currency. The variables of interest are the interest-rate differential (IRD_{USD}, IRD_{EUR}), the nominal exchange rate (FX_{USD}, FX_{EUR}), the VIX index (VIX), 10-year bond yields (Y_{USD}, Y_{EUR}), stock market prices (P_{USD}, P_{EUR}) and carry trade positions (CTF_{USD} and $CTFO_{USD}$, CT_{EUR}). The majority of the data stems from Datastream. A similar set of variables is widely chosen in the literature (see, e.g., Nishigaki, 2007; Brunnermeier et al., 2009 or Ranaldo and Söderlind, 2010).

For the calculation of the IRD_{USD} and IRD_{EUR} we obtain 3-month interbank interest rates. The carries are defined as the difference between the respective target currency interest rate (United States or Euro area) and the Swiss interest rate. Accordingly, we employ the nominal exchange rates CHF/USD, FX_{USD}, as well as CHF/EUR, FX_{EUR}. Furthermore, the VIX volatility index, VIX, from the Chicago Board Options Exchange (CBOE) serves as a proxy for the expected stock market risk.[6]

For an analysis on carry trade positions based on the Swiss and US markets, prices on the US stock exchange market index S&P 500, P_{USD}, and 10-year constant to maturity Treasury bond yields, Y_{USD}, were collected. If the EUR serves as target currency, prices of the euro stock exchange market index Euro Stoxx 50, P_{EUR}, and the synthetic euro benchmark bond yield series,[7] Y_{EUR}, are used.

Trades in the currency markets are usually over-the-counter, making it difficult to find appropriate proxies for carry trade positions. Hence, we rely on data from the U.S.

[6]The index is based on the stock market index S&P 500 and estimates expected volatility by averaging the weighted prices of options over a wide range of strike prices. Brunnermeier et al. (2009) argue that the index is a useful proxy for investor sentiment or "global risk appetite".

[7]The US benchmark bond yield series from Datastream is almost identical to the 10-year constant to maturity Treasury yields for the US market. Hence, the Euro benchmark bond yield series is a reliable proxy for our purposes.

Commodity Futures Trading Commission (CFTC) for carry trade positions with regard to the USD. These contracts are traded on the Chicago Mercantile Exchange (CME). Since October 1992, long and short currency futures positions of non-commercial traders are published periodically. All investors are classified as non-commercial or commercial. Commercial investors have currency risk hedging purposes defined by the CFTC. We are only interested in positions held by those traders who basically trade for speculative purposes.

Burnside et al. (2006) show that a strategy of borrowing the low-interest-rate currency and lending the high-interest-rate currency yields a positive payoff if, and only if, a forward contract has a positive payoff. According to Brunnermeier et al. (2009), few investors actually implement the carry trade using the spot currency market since futures contracts are economically equivalent.[8]

Our proxy for carry trade positions has several shortcomings. First, these data reflect only a very small fraction of currency trades.[9] Second, they are not necessarily results from carry trades, and the classification of commercial and non-commercial traders might be inaccurate in some cases (Galati et al., 2007). Finally, Gagnon and Chaboud (2007) show that the timing of changes in these positions might not be perfectly accurate in all cases. For example, the unwinding of yen carry trades in October 1998 is not displayed in the data.[10] Despite these shortcomings, these futures positions are the best publicly available data (Brunnermeier et al., 2009).

Furthermore, we calculate the so-called "success rate". For the samples considered in our study, we count the observations for which the investors increase the net long futures positions (decrease the net long futures positions) and the CHF appreciates (depreciates) against the USD. The success rate is in the range of 69% and 87%, and above 75% three-quarters of the time. In line with the results of Klitgaard and Weir (2004), we find a strong contemporaneous correlation between changes in net futures positions and exchange rate fluctuations. Thus, knowing the traders actions gives a reasonable chance of correctly estimating the direction of the exchange rate movement during the same week.

A new data set including futures and options was launched from the CME at the end of March 1995. Keeping in mind that an option contract differs in several respects from a futures contract, we use these data for our robustness analysis. From Mogford and Pain (2006) we know that speculative future positions from CME and risk reversals, reflecting the views of options purchasers, move a significant number of times in the same direction.

[8]Futures and forward contracts are similar, yet the former is traded on the stock exchange and the latter over-the-counter. Additionally, they differ in settlement conditions. These differences, however, are not decisive for our purposes.

[9]Following Klitgaard and Weir (2004) a substantial part of the high foreign exchange transaction volume reflects traders' risk management. Hence, the global volume by itself does not preclude the possibility that participations in futures markets might cause currency movements.

[10]The sharp movement to a net long yen position occurred one month before the actual carry trade unwinding (Gagnon and Chaboud, 2007).

Carry trade positions are defined as the difference between short and long futures positions, CTF_{USD}, or as the difference between short and long futures and options positions, $CTFO_{USD}$.

If the net position is positive (negative), investors are involved in carry trades with the CHF as a funding (target) currency. These currency futures position data are not available for the EUR.[11] Thus, we use the carry-to-risk ratio (CTR ratio) as a proxy for carry trade activities, CT_{EUR}. The CTR ratio is defined as the 3-month interest-rate differential divided by the implied volatility derived from 3-month at-the-money exchange rate options.[12] Data on implied exchange rate volatility are taken from Bloomberg.

The choice of the CTR ratio as proxy for carry trade positions has several caveats as the CTR ratio does not represent (carry trade) positions directly. Nevertheless, professional currency market watchers take it as an important indicator for carry trade activities. Furthermore, Galati et al. (2007) find significant correlations between the CTR ratio and futures positions traded at the CME.[13]

We take the natural logarithm of the nominal exchange rates, stock market prices, the VIX index and futures (and options) positions.

3.2 Sample Period and Frequency

The weekly sample period with the USD as target currency starts with 03/28/1995 and ends with 06/24/2008. For our robustness analysis, we estimate the model with different sample lengths. We add observations until the end of 2009 to address the recent financial crisis or start with 10/06/1992.

For model specifications in which the EUR serves as the target currency, we use data for the time period from 01/06/1999 to 06/25/2008.

We determined the data frequency according to the variable with the lowest frequency published, as we expect a strong short-run relationship between the variables included in this study.[14] Futures position data from the CFTC are published weekly, thus leading to a weekly frequency. To ensure comparability along the frequency dimension, we also apply weekly data for the model with the CTR ratio as a proxy for carry trade positions.

[11] Unfortunately, due to data limitations, we are not able to examine further target currencies such as the Australian dollar or the New Zealand dollar.

[12] We limited our analysis to the currency pair CHF/EUR as data on implied exchange rate volatility are not continuously available for other potential target currencies.

[13] These correlations always involve the USD. Moreover, Brunnermeier et al. (2009) argue that the past return of carry trades is perhaps a better measure for carry trade positions than futures data from CME. In this case, the CTR ratio is, owing to its forward-looking nature, also a good proxy in a world with rational market participants.

[14] Brunnermeier et al. (2009) include quarterly data, whereas Nishigaki (2007) estimates his model with monthly data.

4 Methodology

We use a multivariate threshold model to analyze the relationship between key financial and macroeconomic variables focusing on carry trade positions. The choice of the method is based on a descriptive analysis, an econometric test and reported information.

First, the descriptive analysis serves to detect sub-periods separated by an endogenous threshold value of the IRD. The results of this analysis are presented in Figures (1) and (2). The former depicts the 3-month interest-rate differential, IRD_{USD}, between the United States and Switzerland. Until 2001, the IRD_{USD} spread was substantial (about 3% to 4.5%). Subsequently, the difference decreased to around zero percent in November 2001. The following upward trend reaches its maximum of almost 4% at the end of June 2006. The financial crises caused the IRD_{USD} to fall again. Thus, we were able to construct one sub-sample containing large carries and another with smaller differences.[15]

Analogously, Figure (2) illustrates the IRD_{EUR}. The starting point of the sample is the euro launch. The amplitudes of the IRD_{EUR} are not as distinct as for the IRD_{USD}. Nevertheless, three time periods with higher IRD_{EUR} could be identified: the beginning of the sample, the period from mid-2002 to almost the end of 2004 and the end of the sample.

Moreover, these findings are also reflected in the residuals of a regression of the interest-rate differential on a constant and lagged values of all variables. The residuals follow a very similar pattern to the interest-rate differentials themselves.

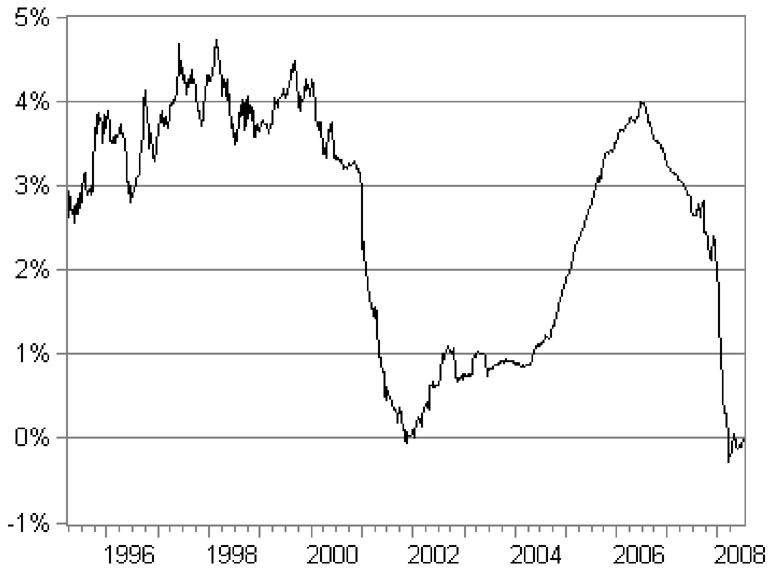


Figure 1: IRD between the US and Swiss 3-month interbank interest rates (IRD_{USD})

[15]Note that we allow the sub-periods to be discontinued, i.e. one sub-period is interrupted by the other one.

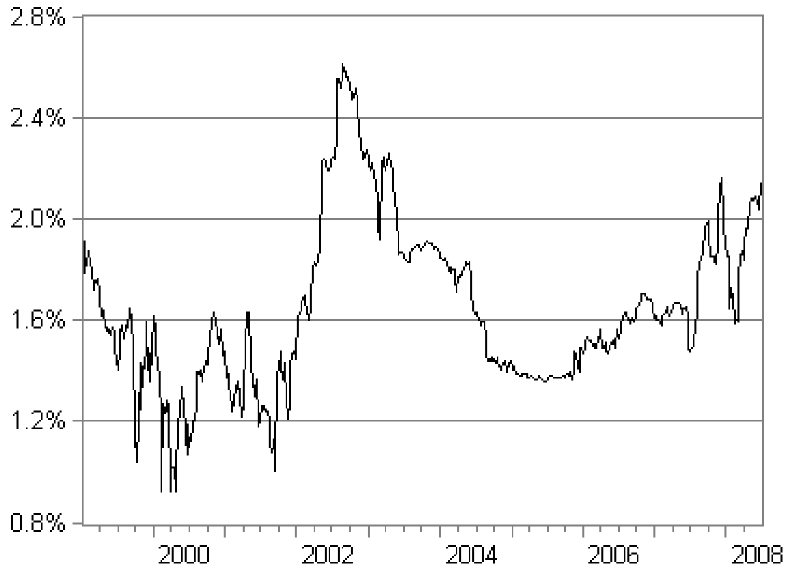


Figure 2: IRD between the Euro and Swiss 3-month interbank interest rates (IRD_{EUR})

Second, the insights of the descriptive analysis are confirmed by the estimation results of a reduced vector autoregressive regression model (VAR) for the whole period. We have to reject the null hypothesis of no autoregressive conditional heteroscedasticity (ARCH) for the majority of error term variances.[16] This is not surprising, since we have high frequency financial variables in our model.[17] Nevertheless, this result indicates a nonlinear relationship between the variables considered.

Finally, professional currency market analysts argue that there exists a threshold level for the carry, above which investor behavior changes.[18] We assume that the dynamic behavior of carry trade positions depends on the magnitude of the carry, and therefore apply a multivariate threshold model for our empirical investigation (Tsay, 1998). Similar methods to study relationships where nonlinear effects are present are used by Canjels et al. (2004), Bernholz and Kugler (2011) and others.

4.1 Multivariate Threshold Model

Before we turn to the econometric model, we test the appropriateness of a multivariate threshold model by applying a test developed by Tsay (1998). The observations are ordered in descending order of the lagged threshold variable to estimate the recursive residuals. The lag is determined by the threshold delay parameter, d. If the dependent

[16]The ARCH test results are summarized in the Tables (10) and (11) in Appendix 1.A.1.

[17]The variance of the error term might follow an ARCH/GARCH process when financial variables are included in a model with high frequency data (see, e.g., Engle, 2001).

[18]I would like to thank the Head FX Research of a major Swiss bank for this important information.

variables are linear, then the recursive least squares estimator of the arranged VAR model is consistent, i.e. the coefficients are zero (Tsay, 1998). Compared to the standard test, we modify its computation to account for conditional heteroscedasticity (Tsay, 1998), i.e. the correlation between the squared error terms and the elements of $X_t'X_t$. The variances of the least squares estimates are adjusted by correcting the weights to standardize the predictive residuals of the recursive least squares estimations. The generalized multivariate threshold model can be written as:

$$\mathbf{y}_t = \mathbf{c}^{(j)} + \Phi_1^{(j)}\mathbf{y}_{t-1} + \cdots + \Phi_p^{(j)}\mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t \quad \text{if } \tau_{j-1} \leq y_{1,t-d} < \tau_j,$$

where $\mathbf{y}_t$ denotes a (6×1) vector containing the values at date t of six variables (interest-rate differential, VIX index, carry trade positions, nominal exchange rate, bond yields, stock market index)[19], $\mathbf{c}^{(j)}$ are the constant vectors for the different regimes, and $\Phi^{(j)}$ denotes a (6×6) coefficient matrix of the respective lag and regime. The vector of error terms is denoted as $\boldsymbol{\epsilon}$, and p is the number of lags included. Let $-\infty = \tau_0 < \tau_1 < \cdots < \tau_{s-1} < \tau_s = \infty$. Then $j = 1, \ldots, s$ represents the different regimes.

We concentrate on models with two regimes, hence, we have only one threshold value and $s = 2$.[20] The multivariate threshold model applied with two regimes has the following form:

$$\mathbf{y}_t = \mathbf{c}^{(1)} + \Phi_1^{(1)}\mathbf{y}_{t-1} + \cdots + \Phi_p^{(1)}\mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t \quad \text{if } y_{1,t-d} < \tau, \tag{1}$$
$$\mathbf{y}_t = \mathbf{c}^{(2)} + \Phi_1^{(2)}\mathbf{y}_{t-1} + \cdots + \Phi_p^{(2)}\mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t \quad \text{if } y_{1,t-d} \geq \tau. \tag{2}$$

The observations of a specific date are included in the first regime (Equation 1) if the threshold variable y_1 is below the threshold value, τ, to the second regime (Equation 2) otherwise. The determination of the delay parameter, d, is based on the test statistic of the Tsay test. In order to determine the threshold value we use a grid search over a reasonable interval of possible values of the threshold variable. The selection of τ is based on the minimized determinant of the variance-covariance matrix. When τ is known, we can estimate the model by ordinary least squares (OLS). Concretely, we estimate the following model:

$$\mathbf{y}_t = \mathbf{c} + (\Phi_1^{(1)}\mathbf{y}_{t-1} + \cdots + \Phi_p^{(1)}\mathbf{y}_{t-p})D_{t-d} + (\Phi_1^{(2)}\mathbf{y}_{t-1} + \cdots + \Phi_p^{(2)}\mathbf{y}_{t-p})(1 - D_{t-d}) + \boldsymbol{\epsilon}_t,$$

where a dummy variable D is defined as being one if $y_{1,t-d} < \tau$, and zero if $y_{1,t-d} \geq \tau$.

[19]The variables enter the model either in level or in first differences. More details on the model specifications can be found in Section 5.

[20]The model was also estimated with two threshold values and with the first difference of the IRD as threshold variable. In these cases, the estimation technique does not change, only the notation becomes slightly more complicated.

4.2 Generalized Impulse Response Functions

Since Sims (1980) seminal paper, vector autoregressions (VARs) are routinely carried out to study dynamic systems. In numerous studies, researchers rely on the Cholesky decomposition to structure the estimation model. Both Nishigaki (2007) and Brunnermeier et al. (2009) use this approach to examine carry trade positions. The structural shocks are obtained by orthogonalizing the estimated reduced-form error terms. However, the ordering of variables in the system matters for the results (Pesaran and Shin, 1998). In many cases it is very difficult to establish a particular recursive ordering on economic theory or institutional knowledge (Stock and Watson, 2001). According to Stock and Watson (2001), researchers are too often tempted to develop economic "theories" that lead to a recursive structure called the "Wold causal chain". Therefore, they distinguish between structural and recursive VARs. Without a widely accepted economic theory to help differentiate between correlation and causation ("identification problem"), we prefer the method developed by Koop et al. (1996) and Pesaran and Shin (1998).[21] This alternative approach is invariant to the ordering of variables, instead, it lacks the possibility of identifying a specific shock.

While the recursive structure identifies the shocks through the Cholesky decomposition of the residual variance-covariance matrix, the variance-covariance matrix itself matters for the generalized impulse response functions. The interdependence of the shocks is carried over to the impulse responses. It follows that the method of generalized impulse response analysis is not the preferred approach for policy statements. In our analysis we do not want to identify specific shocks but rather examine what happens if one of the variables changes unexpectedly. Hence, we let the data speak.

4.3 Confidence Interval: Bootstrap Method

The confidence intervals of impulse responses are routinely computed with bootstrap methods. Kilian (1998b) shows that traditional bootstrap methods such as the frequently applied nonparametric approach developed by Runkle (1987) are inaccurate in the presence of bias and skewness in the small-sample distribution of impulse response estimators. Thus, we adopt his bias-correction (Kilian, 1998b), because the construction of sub-periods reduces the number of observations to a great extent.[22] Additionally, Kilian (1998a) demonstrates the outperformance of the bias-corrected confidence intervals if there is evidence of fat tails or skewness in the error distribution, i.e. the residuals' departure from normality. The distribution of a few estimated residuals in our study suffers

[21] We follow the approach by Pesaran and Shin (1998) as we correct the estimates for small-sample bias and departures from non-normality of the error terms (Kilian, 1998a,b). Furthermore, results from a recursive VAR consistent with Nishigaki (2007) indicate that the GIRFs are reasonable.

[22] Despite the reduction in the number of observations, they are sufficient for an accurate estimation of the parameters.

from non-normality, not only in the full sample but also in the regimes.

As stated earlier, by considering the full samples, we have to reject the null hypothesis of no ARCH effects for the majority of error term variances. However, we conduct the resampling of residuals only within regimes but not across them. The problem is far less severe in the regimes, but it is still present.[23] Non-normality could be at least partly explained by unknown ARCH/GARCH processes.[24] However, as the bias-correction cannot account for biases introduced by ARCH/GARCH processes (Kilian, 1998a), we change the computation of the confidence intervals to deal with unknown ARCH/GARCH processes.

Based on the work by Goncalves and Kilian (2004), we modify the residuals such that we can treat them as *i.i.d.* In order to break up the time interdependence between the estimated residuals we multiply the sequence of residuals with an i.i.d. sequence with mean zero and variance one, drawn from a standard normal distribution. However, we extend the recursive-design wild bootstrap method for univariate models proposed by Goncalves and Kilian (2004) to multivariate models. The application of this method to a multivariate system creates a problem of correctly treating the cross interdependence between residuals of different estimation equations. To overcome this cross interdependence we rely on Pesaran and Shin (1996). In a first step, the residuals are multiplied by the inverse of the Cholesky decomposition:

$$\boldsymbol{\xi} \;=\; \mathbf{P}^{-1}\hat{\boldsymbol{\epsilon}},$$

where $\boldsymbol{\xi}$ is a $(m \times T)$ matrix and $\hat{\boldsymbol{\epsilon}}$ are the estimated residuals. T is the number of observations and m the number of variables. The resulting terms in the matrix $\boldsymbol{\xi}$ are independent from each other for every t. The error terms for which we reject the null hypothesis of no ARCH of order one and/or two and/or four at the 5% significance level are multiplied element by element with i.i.d. sequences described above.[25] The resulting matrix $\boldsymbol{\Gamma}$ has dimension $(m \times T)$. We recover the contemporaneous correlation structure as follows:

$$\hat{\boldsymbol{\epsilon}}^* \;=\; \mathbf{P}\boldsymbol{\Gamma},$$

where $\mathbf{P}$ denotes the Cholesky decomposition matrix. Finally, the matrix $\hat{\boldsymbol{\epsilon}}^*$ contains modified residuals with the same cross interdependence, but no interdependence over time.

[23] Whereas the problem hardly arises in the regime with high interest-rate differentials, it is somewhat stronger in the regime with low interest-rate differentials.

[24] This is true for the leptokurtosis, but not for the skewness in the residuals (Kilian, 1998a).

[25] The computation of the GIRFs requires a constant variance-covariance matrix (Koop et al., 1996). The presence of unknown ARCH/GARCH processes might lead to a time-variant variance-covariance matrix. However, we assume that our results are not strongly biased since we conduct the resampling of residuals only within regimes in which only few or even no error term variances follow an unknown ARCH/GARCH process.

All of these modifications have the property to enlarge the non-centered 95%-confidence intervals of our empirical study. The confidence intervals are based on 11,000 random draws, where the first 1,000 draws are used to compute the bias-correction.[26]

5 Empirical Results

5.1 Preliminary Analysis

In this subsection, we briefly describe the results of the preliminary analysis necessary prior to the estimation of the multivariate threshold model.

5.1.1 Stationarity Tests

In a first step, the time series properties of the variables are examined. For this purpose, the test proposed by Phillips and Perron (1988) and the augmented Dickey and Fuller (1979) unit root test are applied to the variables. Tables (1) and (2) report the results for the models for which the USD serves as the target currency of carry trades. The results point clearly to stationarity of the carry trade positions and the VIX index, regardless of the sample choice. For the 10-year constant to maturity Treasury bond yields the results are borderline. Even if the null hypothesis cannot be rejected, the test statistic is very close to the critical value of the 10% significance level. The remaining three variables, the CHF/USD exchange rate, the price of the S&P 500 and the interest-rate differential are non-stationary.

Table 1: PP and ADF Unit Root Test Results with the USD as Target Currency

	March 1995 - June 2008		March 1995 - Dec 2009	
	PP	ADF	PP	ADF
FX_{USD}	-1.530	-1.539	-1.987	-2.009
P_{USD}	-2.243	-2.178	-2.284	-2.214
VIX	-3.717***	-3.612***	-3.744***	-3.507***
IRD_{USD}	-0.624	-0.720	-0.692	-0.877
Y_{USD}	-3.122	-3.109	-3.420**	-3.385*
Carry Trade Positions				
CTF_{USD}	-6.785***	-6.984***	-7.021***	-7.230***
$CTFO_{USD}$	-6.801***	-6.575***	-7.029***	-7.226***

Notes: FX_{USD}, P_{USD} and Y_{USD}: A deterministic trend is included. PP: Bartlett kernel, Newey-West bandwidth. ADF: Lag length selection by modified SIC (Ng and Perron, 2001). */**/*** denotes significance at 10%, 5% and 1% level, respectively.

[26]Furthermore, if one of the draws leads to a model with an eigenvalue greater than unity (i.e., the model is explosive), the draw is disregarded and repeated.

Table 2: PP and ADF Unit Root Test Results with the USD as Target Currency

	Oct 1992 - June 2008		Oct 1992 - Dec 2009	
	PP	ADF	PP	ADF
FX_{USD}	-1.568	-1.547	-1.946	-1.953
P_{USD}	-1.299	-1.238	-1.292	-1.351
VIX	-3.746***	-3.620***	-3.480***	-3.679***
IRD_{USD}	-2.354	-1.824	-2.200	-2.818
Y_{USD}	-3.197*	-3.043	-3.326*	-3.525**
Carry Trade Positions				
CTF_{USD}	-7.237***	-7.323***	-7.566***	-7.468***

Notes: FX_{USD}, P_{USD}, IRD_{USD} and Y_{USD}: A deterministic trend is included. PP: Bartlett kernel, Newey-West bandwidth. ADF: Lag length selection by modified SIC (Ng and Perron, 2001). */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Table 3: PP and ADF Unit Root Test Results with the EUR as Target Currency

	Jan 1999 - June 2008		Jan 1999 - Dec 2009	
	PP	ADF	PP	ADF
FX_{EUR}	-2.015	-2.107	-1.998	-1.586
P_{EUR}	-1.263	-1.170	-1.519	-1.346
VIX	-2.911**	-2.746*	-2.917**	-2.705*
IRD_{EUR}	-2.098	-2.067	-1.181	-1.027
Y_{EUR}	-1.709	-1.574	-1.732	-1.649
Carry Trade Positions				
CT_{EUR}	-3.461***	-3.603***	-2.127	-1.748

Notes: FX_{EUR}: A deterministic trend is included. PP: Bartlett kernel, Newey-West bandwidth. ADF: Lag length selection by modified SIC (Ng and Perron, 2001). */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Table (3) presents the results for the sample with the EUR as target currency. Again, the proxy for carry trade activities is clearly stationary. The results for the VIX index also points to stationarity. All other time series are non-stationary subject to the test results.

All results are confirmed by applying the Kwiatkowski et al. (1992) stationarity test and the two unit root tests from Elliott et al. (1996) and Ng and Perron (2001). Moreover, all of them point to a (weak) stationary IRD between the 3-month interbank interest rates from Switzerland and the Euro area for the period from January 1999 to June 2008, and a (weak) stationary carry-to-risk ratio for the period from January 1999 to December 2009.[27]

The outcomes of tests for non-stationarity of the time series are in line with the findings of other empirical studies (see, e.g., Nishigaki, 2007). From a theoretical point

[27]These results are not published but can be obtained from the author upon request.

of view it is surprising that the null hypothesis cannot be rejected for the difference between the US and Swiss 3-month interbank interest rates. This result implies that the correct model specification includes the first difference of the IRD. However, there is no economical justification for a random walk behavior of the IRD, specifically in the long run. In addition, the test result might be biased due to the nonlinear threshold nature of this variable. Moreover, as long as the model is stationary and no spurious regression problem arises, the coefficients are estimated consistently, even if the model contains non-stationary variables (Sims et al., 1990). Furthermore, we believe that the divergence of the IRD within the threshold model regimes is much smaller than in the full sample. Hence, the variable might be even stationary.[28] Therefore, we assume that the interest-rate differentials are stationary.[29]

Thus, the model contains the nominal exchange rates (ΔFX_{USD}, ΔFX_{EUR}), the prices of the stock market indices (ΔP_{USD}, ΔP_{EUR}) and ΔY_{EUR} in first differences. The interest-rate differential (IRD_{USD}, IRD_{EUR}), the VIX volatility index (VIX) and the proxies for carry trade activities (CTF_{USD} and $CTFO_{USD}$, CT_{EUR}) enter the model in levels. Furthermore, we assume the 10-year constant to maturity Treasury bond yield series to be trend-stationary and remove the linear trend from the series, Y_{USD}. Following the unit root test results, the series is at least very close to being trend-stationary.[30] Table (4) displays the definitions of the samples.[31] We do not show all results for the samples constructed to analyze the robustness of the findings.[32]

5.1.2 Threshold Nonlinearity Test and Grid Search

Prior to testing threshold nonlinearity, we determine the number of lags included in the model. According to the Akaike & Schwarz lag length selection test results, the optimal lag length is either one or two. But with very few lags included, the estimated residuals exhibit strong serial correlations, as both multivariate and univariate Lagrange multiplier (LM) test results show. Therefore, we must include more lags to avoid endogeneity problems in our estimates. Thus, the choice of the lag length is based on serial correlation tests for the error terms. We tested for serial correlation in the residuals with the multivariate and univariate LM tests of order one, two and four. The optimal lag length of the samples A_{USD}, C_{USD} and E_{USD} is four. For the sample D_{USD}, we choose five, and for

[28] The sample sizes of the sub-periods are too small to get reasonable results from applying unit root tests. This issue is restated in Section 5.4 where the results of the Granger causality tests are discussed.

[29] We also estimated the model with the first difference of the IRD. In contrast to the model with the IRD in levels, we do not find nonlinear effects for all sample periods. For the periods where we do find nonlinear relationships, the results support our findings.

[30] It is well known that these tests have poor power properties relative to the alternative which follows a persistent stationary stochastic process (see, e.g., Christiano et al., 2003)

[31] The subscript to the sample notations indicates the target currency.

[32] These results can be obtained from the author upon request.

Table 4: Sample Definitions

Sample	Period	Variables				
		3-Month LIBOR IRD	Carry Trade Positions	FX	Bond Yields	Stock Market Index
Main Samples						
A_{USD}	March 1995 - June 2008	IRD_{USD}	CTF_{USD}	ΔFX_{USD}	Y_{USD}	ΔP_{USD}
B_{EUR}	Jan 1999 - June 2008	IRD_{EUR}	CT_{EUR}	ΔFX_{EUR}	ΔY_{EUR}	ΔP_{EUR}
Samples for Robustness Analysis						
C_{USD}	March 1995 - Dec 2009	IRD_{USD}	CTF_{USD}	ΔFX_{USD}	Y_{USD}	ΔP_{USD}
D_{USD}	Oct 1992 - June 2008	IRD_{USD}	CTF_{USD}	ΔFX_{USD}	Y_{USD}	ΔP_{USD}
E_{USD}	March 1995 - June 2008	IRD_{USD}	$CTFO_{USD}$	ΔFX_{USD}	Y_{USD}	ΔP_{USD}
F_{EUR}	Jan 1999 - Dec 2009	IRD_{EUR}	CT_{EUR}	ΔFX_{EUR}	ΔY_{EUR}	ΔP_{EUR}

Notes: The sources and more details about the variables are described in Section 3.1. All samples additionally include the VIX index. Y_{USD} is linearly detrended.

sample B_{EUR} and F_{EUR}, two lags.[33] Except for sample B_{EUR}, neither including more lags nor reducing the number of lags improves the serial correlation test results noticeably. We estimate sample B_{EUR} with two instead of three lags, because the threshold model cannot be estimated accurately otherwise.[34] Nevertheless, a few error terms of the models estimated with the optimal lag length still exhibit serial correlation. The test results for the univariate serial correlation LM test are summarized in Table (5). Moreover, the multivariate serial correlation LM test rejects the null hypothesis of no serial correlations of order four for sample A_{USD} at the 5% significance level. For sample B_{EUR}, the null hypothesis of no serial correlations of order one, two and four is rejected at the 10% significance level. The misspecification of a simple linear model might lead to these results.

Table 5: Univariate Serial Correlation LM Test Results

Dependent Variable	Sample A_{USD}			Sample B_{EUR}		
	AR(1)	AR(2)	AR(4)	AR(1)	AR(2)	AR(4)
ΔFX_{USD} / ΔFX_{EUR}	0.040	0.101	3.150	0.081	0.117	5.153
ΔP_{USD} / ΔP_{EUR}	0.026	0.329	5.592	4.828**	6.195*	11.548**
VIX	0.005	0.137	5.006	5.972**	5.966*	7.404
IRD_{USD} / IRD_{EUR}	0.971	2.935	7.269	1.953	6.313**	12.804**
Y_{USD} / ΔY_{EUR}	1.382	2.011	4.522	0.731	2.594	4.238
Carry Trade Positions						
CTF_{USD} / CTF_{EUR}	5.408**	5.335*	5.598	1.737	5.638*	6.131

Notes: The samples are described in Table (4). The LM test results are based on four lags for sample A_{USD} and two lags for sample B_{EUR}. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

The Tsay test to detect threshold nonlinearity, corrected for the possibility of conditional heteroscedasticity, is applied with delay parameters, d, equal to one, two and three.[35] For reasons discussed in Section 4, we choose the interest-rate differential as the threshold variable. The findings for all samples are shown in Table (6). Overall, we conclude that for the majority of model specifications we can reject the null hypothesis of parameter stability. If threshold nonlinearity is present for more than one value of d, we aim to choose d such that it corresponds to the maximum of the Chi-squared test statistic. For different reasons this is not always achievable. The threshold value for sample B_{EUR} with $d = 2$ leaves for one of the two regimes too few observations for an accurate estimation. Hence, we set the delay parameter equal to three. Sample F_{EUR} is estimated with $d = 1$ because one of the regimes has an eigenvalue greater than unity with $d > 1$. For sample A_{USD} we choose $d = 3$ instead of $d = 2$, because the latter value is preferred

[33] The results of sample D_{USD} are robust to the estimation with four lags.

[34] The threshold value determined to detect the two regimes leaves for one regime too few observations for reliable estimations.

[35] More details on the applied Tsay test are provided in Section 4.1.

for the samples C_{USD} and D_{USD}.[36] Sample E_{USD} is estimated with the delay parameter equal to three for purposes of comparison. As the differences between the test statistics are small, sample A_{USD} is estimated with $d = 1$ and $d = 3$ to check for possible variations in the impulse response functions. Our main model specifications are $A^{d=3}_{USD}$ and $B^{d=3}_{EUR}$. All versions estimated are denoted by extra bold type.

Table 6: Results of the Tsay Test

Sample	Delay Parameter (d)		
	1	2	3
Main Samples			
A_{USD}	**221.2**(150)***	212.1(150)***	**211.5**(150)***
B_{EUR}	219.6(78)***	274.1(78)***	**273.5**(78)***
Samples for Robustness Analysis			
C_{USD}	225.4(150)***	190.0(150)**	**247.4**(150)***
D_{USD}	238.7(150)***	206.0(150)***	**222.2**(150)***
E_{USD}	214.6(150)***	220.5(150)***	**183.3**(150)**
F_{EUR}	**160.6**(78)***	204.3(78)***	187.6(78)***

Notes: The samples are described in Table (4). The estimated models are denoted by extra bold type. The degrees of freedom are written in brackets. */**/*** denotes significance of the Chi-squared value at 10%, 5% and 1% level, respectively.

In order to estimate the multivariate threshold model, the threshold values for all model specifications are determined. The selection of the threshold value, τ, is based on a grid search for the minimized determinant of the variance-covariance matrix. Table (7) depicts τ for the different models. As shown in Figures (1) and (2), IRD_{USD} and IRD_{EUR} are almost always positive over all sample periods. Therefore, we search for a value which separates two regimes depending on the size of the carry. One regime contains observations with values of the threshold variable greater than or equal to τ, all other observations are collected in the other regime. The threshold values are between 1.84% and 2.94%. Compared to $A^{d=3}_{USD}$, τ falls if additional observations until the end of 2009 are added (sample C_{USD}) or if a smaller delay parameter value is chosen ($d = 1$). The contrary is true for sample D_{USD} starting with 10/06/1992. The inclusion of options positions does not alter the result.[37]

[36] The eigenvalue of one regime of the model is greater than unity when sample D_{USD} is estimated with $d = 1$.

[37] In addition, for the main samples, we searched for two threshold values instead of one. The minimized determinant of the variance-covariance matrix of sample A_{USD} increases in the specification with two threshold values. Therefore, the model with one threshold value is preferred. For sample B_{EUR} the minimized determinant is smaller. However, as the grid search reveals that one threshold value is almost equal to 1.84% and the other is very close to the minimum value of IRD_{EUR}, we consider only models with one threshold value.

Table 7: Threshold Values (Percentage)

Sample	Delay Parameter (d)	
	1	3
Main Samples		
A_{USD}	2.12	2.63
B_{EUR}		1.84
Samples for Robustness Analysis		
C_{USD}		2.17
D_{USD}		2.94
E_{USD}		2.63
F_{EUR}		1.91

Notes: The samples are described in Table (4).

5.2 Estimated Generalized Impulse Responses

In this section, we discuss the generalized impulse response functions (GIRFs) of the main samples $A^{d=3}_{USD}$ and $B^{d=3}_{EUR}$. For sample $A^{d=3}_{USD}$ we compute the GIRFs for the regime with values of $IRD^{d=3}_{USD}$ greater than or equal to the threshold value of 2.63%. This regime is denoted as H-regime. The GIRFs for the regime with values of $IRD^{d=3}_{USD}$ smaller than 2.63% are shown in the L-regime. The same approach determines the GIRFs of sample $B^{d=3}_{EUR}$ with the threshold variable $IRD^{d=3}_{EUR}$ and the threshold value of 1.84%.[38] We present the point estimates (solid line), the median of the bootstraps (dashed-dotted line) and the non-centered 95%-confidence interval (dotted lines).[39]

Figure (3) shows the (accumulated) GIRFs of the sample $A^{d=3}_{USD}$ VAR system in response to a one-standard deviation IRD_{USD} shock in the H-regime. An unexpected increase in IRD_{USD}, through an increase in the US interest rate and/or a decrease in the Swiss interest rate, leads to a statistically significant contemporaneous rise in VIX, a decline in CTF_{USD} and P_{USD}, as well as an appreciation of the Swiss currency. The impacts on CTF_{USD} and P_{USD} last slightly longer than one week. While the increased IRD_{USD} improves the environment for a profitable carry trade strategy, other variables such as risk sentiment and US stock market prices indicate a rising risk for a sudden and strong unwinding of carry trades. This result echoes the finding of Brunnermeier et al. (2009) that the conditional skewness becomes more negative after an interest-rate differential shock. The response of FX_{USD} is partially in line with the prediction of uncovered interest rate parity (UIP). The immediate appreciation of the low-interest-rate currency could be affected by the fall in CTF_{USD}, among other factors such as the decrease in

[38]The (accumulated) GIRFs of all variables at a forecast horizon up to 40 weeks are summarized in Appendix 1.A.2, Tables (12)-(15).

[39]More information on the bootstrap method used to determine the confidence interval is given in Section 4.3.

the investors risk appetite. The so-called "safe haven" property of the CHF might be an explanation for the lack of the initial USD appreciation. Clarida et al. (2009) show that in high exchange rate volatility environments the low-yielding currency tends to appreciate even more than implied by UIP.

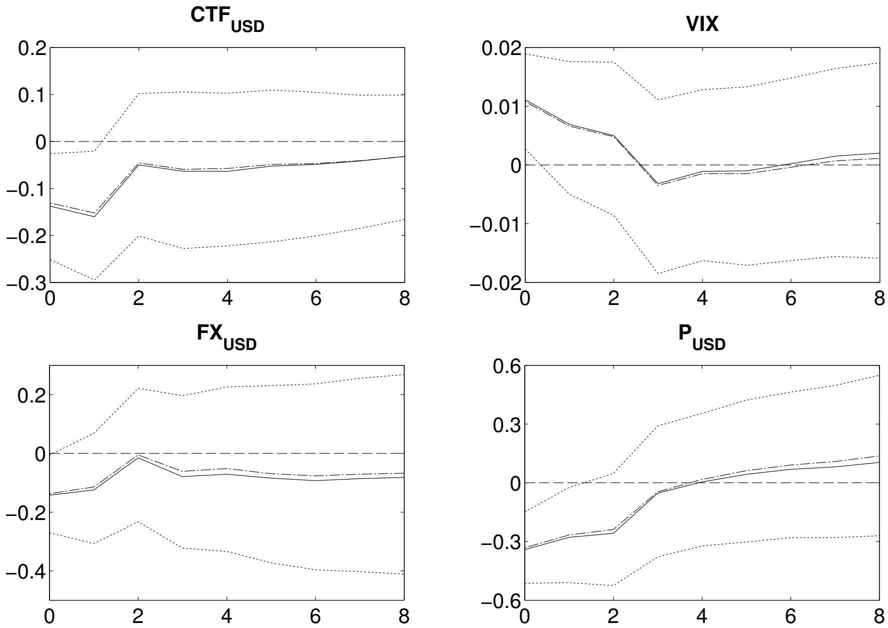


Figure 3: Sample $A^{d=3}_{USD}$: (Accumulated) GIRFs of the variables CTF_{USD}, VIX, ΔFX_{USD} and ΔP_{USD} in response to a one-standard deviation ΔIRD_{USD} shock in the H-regime. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $A^{d=3}_{USD}$ see Table (4). Number of observations: 418

In the L-regime the effects are different (see Figure 4). In the short run none of the responses are statistically significant. Nevertheless, some long-run trends are revealed. The shock tends to result in a lower risk sentiment, a continuous depreciation of the CHF, pointing to the UIP puzzle, and an increase in P_{USD}. Although in the short run CTF_{USD} hardly moves, in the period between five and ten months after the shock, the buildup of CTF_{USD} is statistically significant. However, the *insignificant* appreciation of the USD on impact and its trend to further appreciate instead of a CHF appreciation as UIP predicts, could be due to the under reaction of carry trade activities. Brunnermeier et al. (2009), who do not distinguish between different interest-rate differential regimes, infer that carry trade activities in response to a shock is not enough to push up the exchange rate towards the value implied by UIP. To summarize, in the H-regime a further increase in the carry leads rather to a fall in CTF_{USD} in the short run and in the L-regime to a rise in the long run. These opposed effects arise due to different risk environments, liquidity constraints and/or exchange rate fluctuations. While the carry is the key variable determining carry trades in the model of Plantin and Shin (2010), our result suggests that the associated changes in risk and the exchange rate are empirically significant.

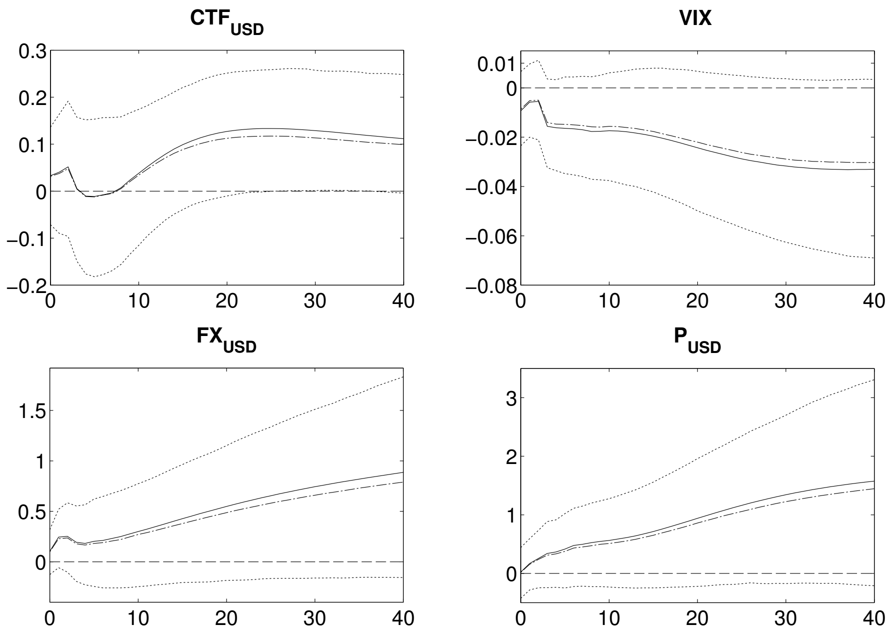


Figure 4: Sample $A^{d=3}_{USD}$: (Accumulated) GIRFs of the variables CTF_{USD}, VIX, ΔFX_{USD} and ΔP_{USD} in response to a one-standard deviation ΔIRD_{USD} shock in the L-regime. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $A^{d=3}_{USD}$ see Table (4). Number of observations: 270

A simple analysis of a sudden and strong movement of ΔFX_{USD}, approximated by $1.64\sigma_{\Delta FX_{USD}}$ and $1.96\sigma_{\Delta FX_{USD}}$, reveals that in the H-regime a strong appreciation of the Swiss currency happens twice as often as a strong depreciation, while in the L-regime the fraction is 52% and 57%, respectively. This finding mirrors the results obtained by Brunnermeier et al. (2009). The authors conclude that in times when the IRD is high, the skewness of carry trade returns is particularly negative. The higher probability of a sudden appreciation ("crash") of the Swiss Franc in the H-regime might be attributed to differences in fundamentals. The average monthly CPI inflation-rate differential between the US and Switzerland is in the H-regime 0.5 percentage points higher than in the L-regime (1.94% vs. 1.40%).

A shock to VIX gives rise to a statistically significant contraction of CTF_{USD} in both regimes, shown in Figure (5). This pattern is not surprising, as an increase in VIX represents a higher risk sentiment and it is in line with the results found by Nishigaki (2007) and Brunnermeier et al. (2009). The decline is slightly stronger in the H-regime (left panel), reflecting an increased risk aversion of the speculators with a larger carry. The effects on FX_{USD} and P_{USD}[40] are similar across both regimes. Yet, the initial decrease in FX_{USD} is somewhat larger in the L-regime (right panel).

What happens to the variables in the VAR after an unexpected unwinding of carry trades? Brunnermeier et al. (2009), for instance, conjecture that sudden exchange rate

[40]See Figures (12) and (13) in Appendix 1.A.2.

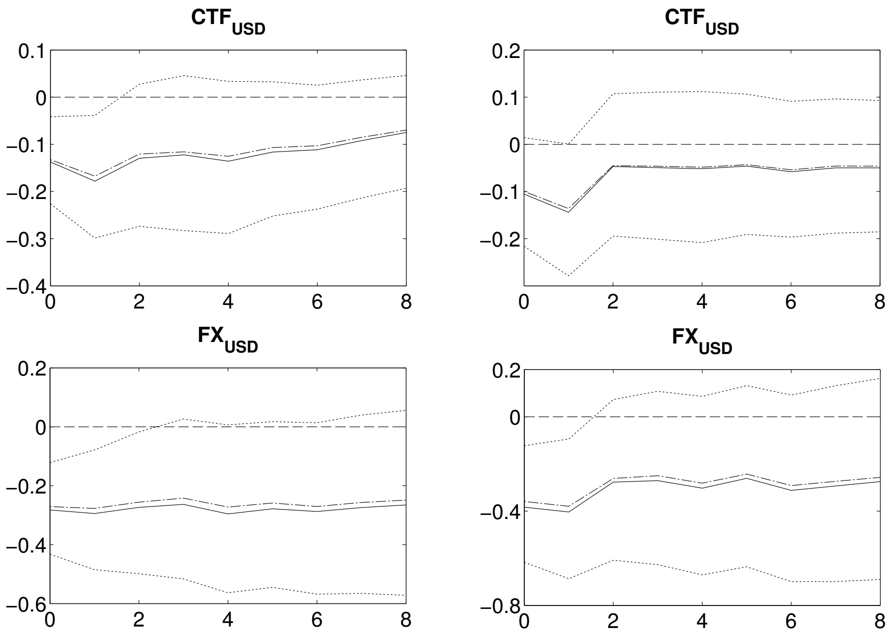


Figure 5: Sample $A^{d=3}_{USD}$: (Accumulated) GIRFs of the variables CTF_{USD} and ΔFX_{USD} in response to a one-standard deviation VIX shock. The left panel depicts the H-regime, the right panel depicts the L-regime. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $A^{d=3}_{USD}$ see Table (4). Number of observations: 418 (H-regime) & 270 (L-regime)

fluctuations unrelated to fundamental news events can be triggered when investors near funding constraints. We expect a strong appreciation of the CHF as the demand for the Swiss currency rises sharply. Figure (6) confirms this assumption. The currency appreciates contemporaneously in both regimes.[41] A shock whose size is twice the standard deviation of CTF_{USD} leads to an immediate appreciation of the CHF of about three percent in the H-regime (left panel) and four percent in the L-regime (right panel). In the L-regime the CHF starts to depreciate after a sudden appreciation. The effect diminishes over time and ceases to be statistically significant after four months (see Figure 13 in Appendix 1.A.2). In contrast, we find a slight overshooting in the H-regime, and the Swiss currency remains appreciated against the US currency over the entire forecast horizon. In the study of Nishigaki (2007), the appreciation of the yen is also statistically significant and lasts for almost two years. Additionally, in both regimes we find an increase in VIX. Whereas in the H-regime the effect is statistically significant in the short run, in the other regime it is significant in the medium run too.

Figure (7) shows that an unexpected depreciation of the Swiss franc results in a large and statistically significant buildup of CTF_{USD}. The reduction of the positions over time

[41]The variance decomposition based on the Cholesky decomposition ordering in line with Nishigaki (2007); IRD_{USD}, VIX, CTF_{USD}, ΔFX_{USD}, Y_{USD} and ΔP_{USD}, reveals that the semi-structured carry trade activities shock explains about 25% of FX_{USD} in both regimes. It is the most important shock apart from the own shock.

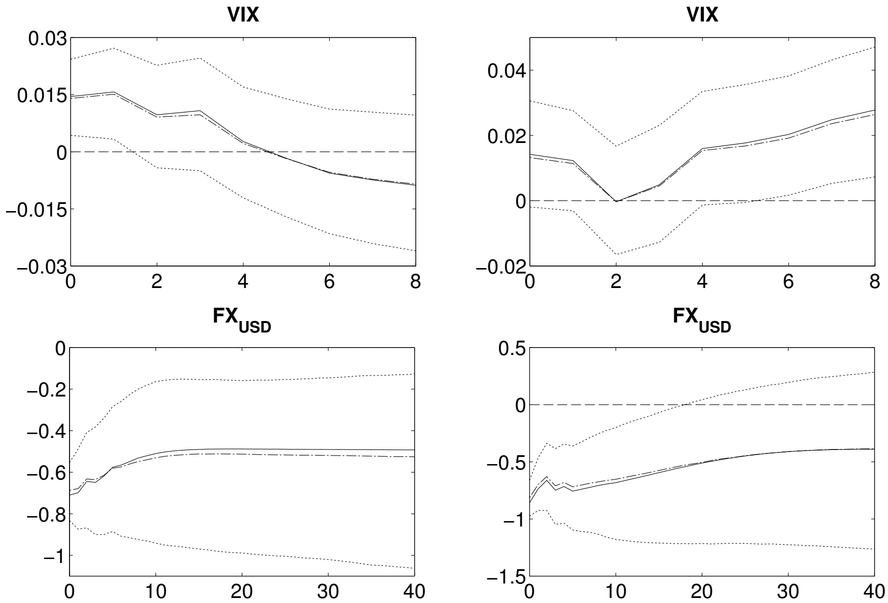


Figure 6: Sample $A^{d=3}_{USD}$: (Accumulated) GIRFs of the variables VIX and ΔFX_{USD} in response to a one-standard deviation CTF_{USD} shock. The left panel depicts the H-regime, the right panel depicts the L-regime. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $A^{d=3}_{USD}$ see Table (4). Number of observations: 418 (H-regime) & 270 (L-regime)

is (marginally) slower in the L-regime (right panel). This could be due to the slower mean reversion of VIX, which falls after the shock in both regimes, and the longer statistically significant increase in Y_{USD} in the L-regime.

An unexpected rise in P_{USD} induces a sudden drop in VIX and an appreciation of the US currency in both regimes (Figure 8). Both effects last longer in the H-regime (left panel). This might be an explanation for the longer horizon over which CTF_{USD} increases, although not statistically significant for all horizons (see Figure 12 in Appendix 1.A.2). Positive shocks to P_{USD} increase the value of a stock portfolio investors would like to use as collateral for liquidity, to engage in carry trade activities. Nishigaki (2007) finds a persistent fall in yen carry trade positions after a negative US stock market shock.

Now we turn to the results for sample $B^{d=3}_{EUR}$. Not surprisingly, a positive innovation to IRD_{EUR} results in a statistically significant rise in CT_{EUR}, which has IRD_{EUR} as its numerator (Figure 9).[42] However, compared to IRD_{EUR} the rise is smaller, hence, the implicit nominal exchange rate volatility increases too. In the long run, depicted in Figure (14) in Appendix 1.A.2, the effect on CT_{EUR} is statistically significant for a longer time span in the L-regime. Apart from the fact that the increase in IRD_{EUR} is statistically significant for a longer period, the negative trend of VIX, and the increase in FX_{EUR}, Y_{EUR} and P_{EUR} might influence this pattern (see Figure 14 in Appendix 1.A.2). This finding is comparable to the results for sample $A^{d=3}_{USD}$.

[42]The correlation between IRD_{EUR} and CT_{EUR} amounts to 0.5.

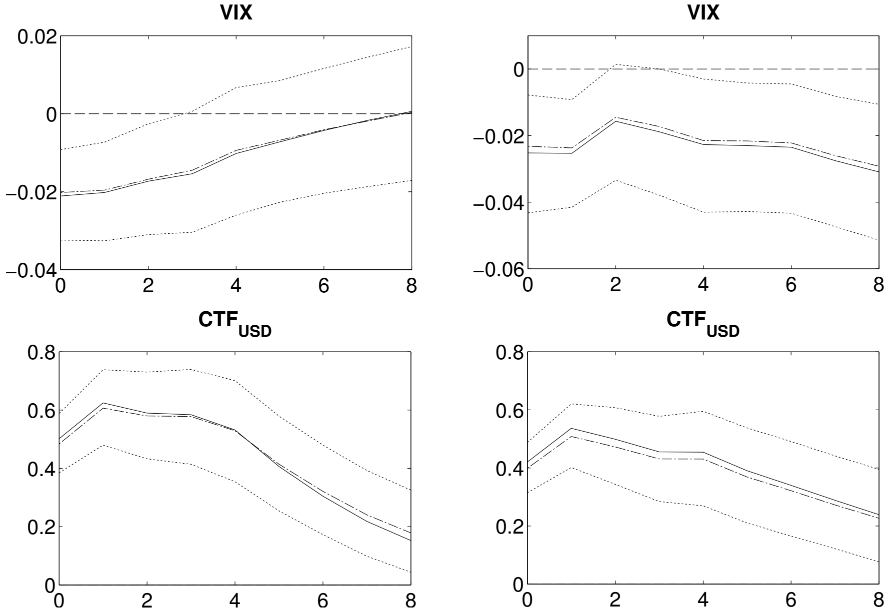


Figure 7: Sample $A^{d=3}_{USD}$: (Accumulated) GIRFs of the variables VIX and CTF_{USD} in response to a one-standard deviation ΔFX_{USD} shock. The left panel depicts the H-regime, the right panel depicts the L-regime. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $A^{d=3}_{USD}$ see Table (4). Number of observations: 418 (H-regime) & 270 (L-regime)

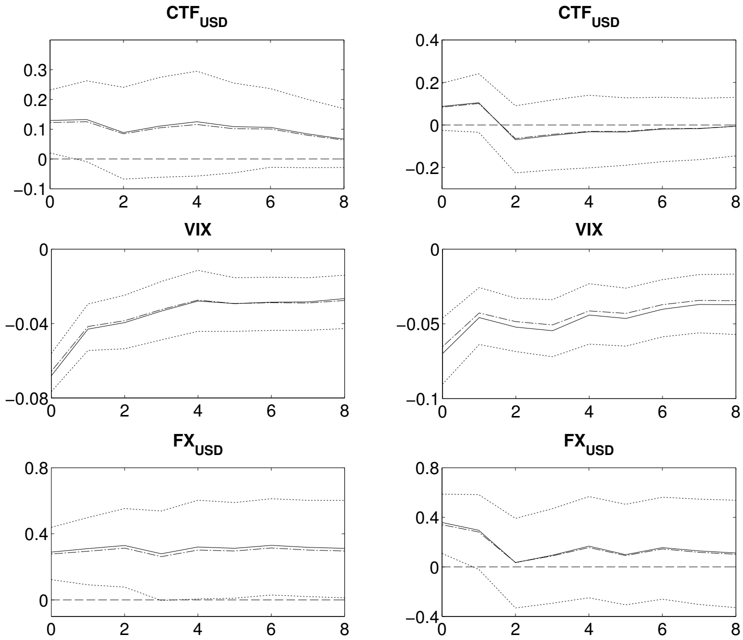


Figure 8: Sample $A^{d=3}_{USD}$: (Accumulated) GIRFs of the variables VIX, CTF_{USD} and ΔFX_{USD} in response to a one-standard deviation ΔP_{USD} shock. The left panel depicts the H-regime, the right panel depicts the L-regime. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $A^{d=3}_{USD}$ see Table (4). Number of observations: 418 (H-regime) & 270 (L-regime)

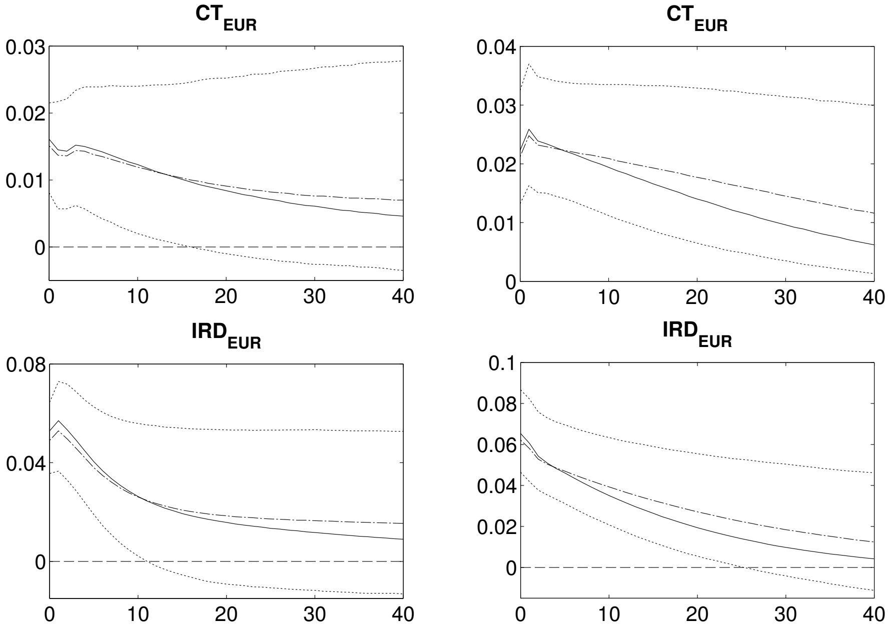


Figure 9: Sample $B^{d=3}_{EUR}$: (Accumulated) GIRFs of the variables CT_{EUR} and ΔIRD_{EUR} in response to a one-standard deviation ΔIRD_{EUR} shock. The left panel depicts the H-regime, the right panel depicts the L-regime. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $B^{d=3}_{EUR}$ see Table (4). Number of observations: 125 (H-regime) & 367 (L-regime)

The analysis of the exchange rate exhibits that strong appreciations of the CHF, approximated by $1.64\sigma_{\Delta FX_{EUR}}$ and $1.96\sigma_{\Delta FX_{EUR}}$, occur more probable in the H-regime than in the L-regime. With equal probability ΔFX_{EUR} should fall in one-quarter of the cases in the H-regime, determined by the number of observations. Yet, 32% ($1.64\sigma_{\Delta FX_{EUR}}$) and 44% ($1.96\sigma_{\Delta FX_{EUR}}$) of the appreciations happen in the H-regime.[43] However, the average monthly CPI inflation-rate differential between the Euro zone and Switzerland is only very slightly higher in the H-regime compared to the L-regime.[44]

Compared to sample $A^{d=3}_{USD}$, the GIRFs associated with an innovation to VIX are qualitatively similar. However, Figure (10) displays that the effects are more pronounced in the L-regime (right panel). The fall in CT_{EUR} in the L-regime could be driven by the strong appreciation of the CHF against the EUR. By virtue of the faster mean reversions of VIX and FX_{EUR} in the H-regime than in the L-regime, CT_{EUR} rises in the long run (see Figure 14 in Appendix 1.A.2).

A one-standard deviation shock to CT_{EUR} gives rise to an expected appreciation of the CHF.[45] Figure (11) shows that the initial impact is equal for both regimes, but

[43]In contrast to sample $A^{d=3}_{USD}$ we do not find a higher probability for a strong appreciation compared to a strong depreciation. The probability is about equal.

[44]The results are sensitive to some negative and positive outliers of the inflation-rate differential.

[45]The variance decomposition based on the Cholesky decomposition ordering in line with Nishigaki (2007); IRD_{EUR}, VIX, CT_{EUR}, ΔFX_{EUR}, ΔY_{EUR} and ΔP_{EUR}, reveals that the semi-structured carry trade activities shock explains about 5% of FX_{EUR} in the H-regime. Apart from the own shock it is the

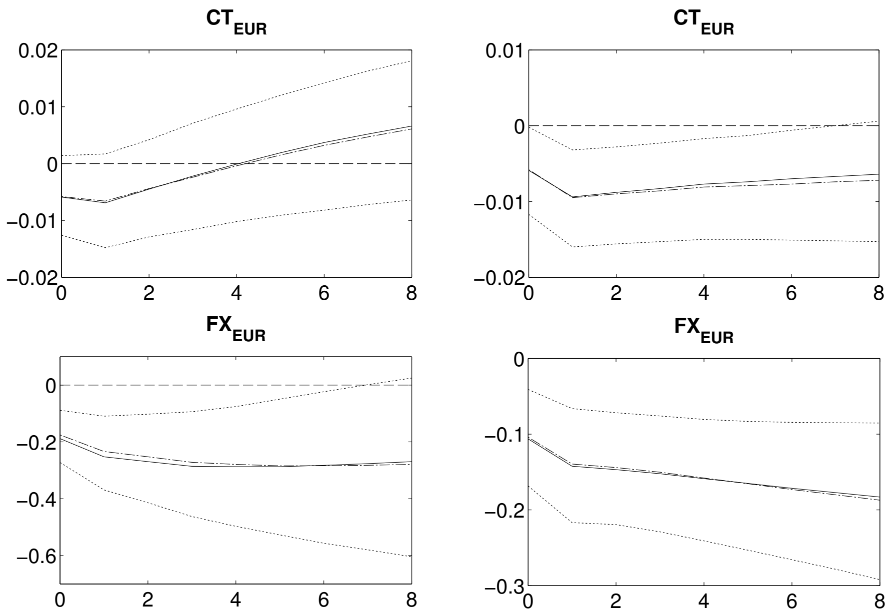


Figure 10: Sample $\mathrm{B}_{EUR}^{d=3}$: (Accumulated) GIRFs of the variables CT_{EUR} and ΔFX_{EUR} in response to a one-standard deviation VIX shock. The left panel depicts the H-regime, the right panel depicts the L-regime. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $\mathrm{B}_{EUR}^{d=3}$ see Table (4). Number of observations: 125 (H-regime) & 367 (L-regime)

the mean reversion is slower in the L-regime (right panel). If the shock equals twice the standard deviation of CT_{EUR} the sudden appreciation of the CHF is slightly more than one percent in both regimes. This effect is smaller compared to the sample $\mathrm{A}_{USD}^{d=3}$. Though, as the proxy for carry trade activities differs, a one-to-one comparison is impossible. Additionally, we find an increase in VIX in the short run and a fall in P_{EUR}. However, the impacts are only significant in the L-regime (Figures 14 and 15 in Appendix 1.A.2).

In line with sample $\mathrm{A}_{USD}^{d=3}$, an unexpected depreciation of the Swiss currency leads to a fall in risk sentiment and an increase in CT_{EUR} and P_{EUR}. Furthermore, the short-run effects of a shock to ΔP_{EUR} are qualitatively the same as in sample $\mathrm{A}_{USD}^{d=3}$, except for Y_{EUR} in the H-regime. In the L-regime the rise in CT_{EUR} becomes marginally statistically significant two weeks after the shock. The stronger impact compared to the H-regime may be a consequence of the severe and persistent depreciation of the Swiss currency.[46]

Overall, we note that there are substantial differences across regimes depending on the size of the carry. Furthermore, the comparison of the two samples reveals that risk sentiment, exchange rates, bond yields and stock market indices show similar (qualitative) patterns with few exceptions, especially for the exchange rate and bond yields. Carry traders seem to react likewise, although the proxies for carry trade activities differ.

second most important shock. In the L-regime it is the most important shock apart from the own shock and explains about 16% of FX_{EUR}.

[46]For the last two shocks compare Figures (14) and (15) in Appendix 1.A.2.

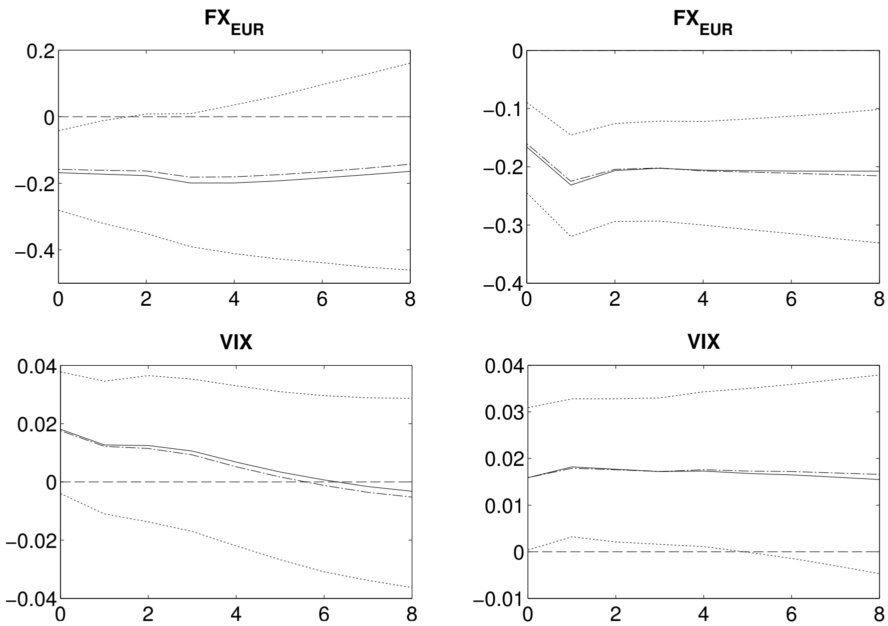


Figure 11: Sample $B_{EUR}^{d=3}$: (Accumulated) GIRFs of the variables ΔFX_{EUR} and VIX in response to a one-standard deviation CT_{EUR} shock. The left panel depicts the H-regime, the right panel depicts the L-regime. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $B_{EUR}^{d=3}$ see Table (4). Number of observations: 125 (H-regime) & 367 (L-regime)

5.3 GIRFs: Robustness Analysis

Overall, the robustness analysis demonstrates robust findings across the different samples. In the following, we describe the changes and point out some important qualitative and quantitative divergences from sample $A_{USD}^{d=3}$ and sample $B_{EUR}^{d=3}$.

5.3.1 Delay Parameter

Since the Chi-squared test statistic for the delay parameter equal to one is the largest among the different delay parameters (see Table 6), we also estimate sample $A_{USD}^{d=1}$. While the GIRFs of the H-regime reveal no qualitative or quantitative differences, the positive long-run impact of a shock to IRD_{USD} on CTF_{USD} is not statistically significant in the L-regime. This might be due to the somewhat faster mean reversion of the carry, a slightly smaller decline in VIX and a less pronounced depreciation of the Swiss currency.

5.3.2 Sample Period Selection

We extended the sample period to include observations of the recent financial crises (sample $C_{USD}^{d=3}$). The GIRFs of the H-regime are very robust to this modification. Yet, several GIRFs of the L-regime exhibit distinct differences compared to the results of sample $A_{USD}^{d=3}$. A one-standard deviation shock to IRD_{USD} has no impact on VIX, FX_{USD} or P_{USD} anymore, i.e. the responses show no trend either way. The absence of these trends might explain that investors do not increase CTF_{USD} in the long run. Besides the modification of the sample length, the reduction in the threshold value determines this result (see Table 7). The mean reversion of FX_{USD} after an unexpected unwinding of carry trades takes longer in sample $\mathrm{C}_{USD}^{d=3}$. Moreover, the impacts on VIX and P_{USD} are no longer statistically significant. This also holds when the Swiss currency depreciates unexpectedly. In general, the confidence intervals for the impulse response functions for the L-regime are expanded, pointing to increased uncertainty during the financial crisis.

The same modification for sample $\mathrm{B}_{EUR}^{d=3}$ reveals that the results of the H-regime are qualitatively robust (sample $\mathrm{F}_{EUR}^{d=1}$).[47] While in sample $\mathrm{F}_{EUR}^{d=1}$ CTF_{EUR} asymptotes faster to its steady-state level after an IRD_{EUR} shock, the effect on carry trades is more persistent in response to an unexpected depreciation of the CHF. Moreover, the decline in CTF_{EUR} becomes statistically significant after a sudden increase in VIX. The same is true for the contemporaneous rise in VIX to a CTF_{EUR} shock. In distinction from sample $\mathrm{A}_{USD}^{d=3}$, the GIRFs of the L-regime do not change markedly by extending the sample period. Nevertheless, compared to the benchmark, the mean reversion of CTF_{EUR} is notably slower after a surprising increase in risk sentiment, the Euro Stoxx index and the exchange rate in sample $\mathrm{F}_{EUR}^{d=1}$.[48] Additionally, the exchange rate remains significantly below its steady-state level for 17 weeks in response to an unwinding of carry trades. This is about two months less than in sample $\mathrm{B}_{EUR}^{d=3}$.

Because weekly published CME futures positions are available since October 1992, sample $\mathrm{D}_{USD}^{d=3}$ contains data from 1992/10/06 until 2008/06/24. The GIRFs of the H-regime are robust to this modification. In contrast to sample $\mathrm{A}_{USD}^{d=3}$, the rise in VIX in response to an unexpected decrease in CTF_{USD} only marginally fails to pass the 5% significance level. However, more substantial changes are observed for the L-regime. A shock to IRD_{USD} gives rise to a statistically insignificant increase in CTF_{USD} in the medium run. This lack of significance is somewhat surprising, because the fall in VIX is statistically significant during three weeks. Yet, after an initial tendency to depreciate, the Swiss currency does not continue to follow a depreciation trend in the long run. An unexpected rise in VIX leads to a longer appreciation of the CHF and fall in P_{USD}. Furthermore, the decline in CTF_{USD} is less pronounced and far from being statistically significant. In the medium run, VIX, Y_{USD} and P_{USD} cease to respond statistically significantly to a sudden unwinding of carry trades. Moreover, FX_{USD} exhibits a slower mean reversion. Finally, when P_{USD} goes up unexpectedly, the increase in CTF_{USD} is statistically significant on impact, in contrast to the jump in sample $\mathrm{A}_{USD}^{d=3}$. This change arises due to the increase in the threshold value, whereas all the other deviations cannot be ascribed to a threshold value change.

[47]Note that the delay parameter d shifts in addition to the change in the sample period.

[48]These changes partly depend on the increased value of the delay parameter (see Table 7).

5.3.3 Futures and Options Positions

The inclusion of options to the CME futures positions to proxy carry trade activities (sample $\mathrm{E}_{USD}^{d=3}$) causes no qualitative change in either regime. However, for the L-regime, the decline in $CTFO_{USD}$ in response to an innovation to VIX is statistically significant for the first week. The same is true for the rise after a shock to P_{USD}. Furthermore, an unexpected increase in Y_{USD} has a slightly longer statistically significant impact on $CTFO_{USD}$. In the H-regime the reaction of $CTFO_{USD}$ to an unexpected increase in IRD_{USD} is slightly less pronounced.

5.3.4 Choice of the Interest Rate

Next, we assess whether the chosen interest rate has any impact on our results. Therefore, we replace the 3-month interbank interest rates with the 1-month interbank interest rates. While this replacement of the interest-rate differentials has no impact on the GIRFs with the USD as target currency, the rise in CTF_{EUR} in response to a sudden increase in IRD_{EUR} is no longer statistically significant in the H-regime.

5.4 Granger Causality Analysis

In this section, we shed light on the question of whether one variable in our models moves ahead of the others, i.e. if the variables "Granger-cause" each other. Following Klitgaard and Weir (2004) and Mogford and Pain (2006), position data do not help in anticipating exchange rate movements for the subsequent week. Their insights are based on a Granger (1969) causality test with two variables, the net futures positions and the nominal exchange rate.[49] We extend their analysis in two ways. First, we include additional variables in our model which have the potential to "Granger-cause" another variable. Second and more important, we distinguish the effects between regimes, depending on the size of the interest-rate differential (IRD). If the value of the threshold variable is greater than or equal to the threshold value, the corresponding observations are assigned to the H-regime, to the L-regime otherwise.[50]

In a first step, the proxy for carry trade positions is excluded from the multivariate threshold model to examine the power of this variable to "Granger-cause" the other variables in the model. Table (8) displays the findings of all samples for each regime.[51]

[49] In their studies, both variables are first differenced prior to the estimation.

[50] Tables (12) and (13) in Appendix 1.A.1 show all results for the main samples.

[51] If a VAR model contains one or more random walk series without cointegration relationship, the Granger causality test statistics have a nonstandard limiting distribution (Sims et al., 1990). The unit root tests reveal that the variable IRD is non-stationary. Nevertheless, we assume this series to be stationary and refer to the standard test statistics since the spread of the IRD is smaller within the regimes compared to the full sample. Further, there is no economic reason for a random walk behavior. The sample sizes of the regimes are too small to get reasonable results from applying unit root tests.

Table 8: Granger Causality Test: Carry Trade Positions Excluded

Sample / Regime	Variable				
	IRD_{USD} / IRD_{EUR}	VIX	ΔFX_{USD} / ΔFX_{EUR}	Y_{USD} / ΔY_{EUR}	ΔP_{USD} / ΔP_{EUR}
$A^{d=3}_{USD}$					
H	10.560**	9.865**	2.444	3.758	6.493
L	1.390	17.172***	10.206**	14.241***	16.136***
$B^{d=3}_{EUR}$					
H	1.575	0.257	0.373	7.325**	0.617
L	22.717***	1.862	8.281**	7.832**	11.539***
$C^{d=3}_{USD}$					
H	14.376***	5.547	2.563	4.288	4.806
L	2.610	10.271**	2.945	2.015	7.388
$D^{d=3}_{USD}$					
H	12.945**	8.783*	2.365	5.058	6.532
L	2.287	8.363*	6.344	9.874**	4.577
$E^{d=3}_{USD}$					
H	2.894	6.961	4.979	2.072	4.365
L	1.290	15.622***	7.904*	15.800***	14.855***

Notes: The samples and variables are described in Table (4). Observations for which the threshold variable lies above the threshold value are assigned to the H-regime; for values below the threshold values, the observations are included in the L-regime. The threshold values are given in Table (7). */**/*** denotes significance of the Chi-squared value at 10%, 5% and 1% level, respectively.

In all three models containing futures position data as proxy for carry trade positions, these positions have predictive power for the IRD in the H-regime. The contrary is true for sample $B^{d=3}_{EUR}$, where carry trade activities lead the IRD in the L-regime. This highly statistically significant result, however, has to be interpreted with caution as the IRD is the numerator of the carry-to-risk ratio (CTR ratio), which is the proxy for carry trade positions.

However, the predictive power of the proxy for carry trade activities is often statistically (more) significant in the L-regime, for example, with respect to nominal exchange rate fluctuations. For all samples, the Chi-squared values for the L-regime are substantially larger, and in addition, in two cases statistically significant at the 5% level and once at the 10% level. This result challenges the insights of Klitgaard and Weir (2004) and Mogford and Pain (2006) in the sense that in times with small IRDs there is the possibility that past position data help to predict exchange rate movements. The knowledge about speculative future positions seems to have incremental information about future fluctuations in the exchange rate in line with findings from the literature, pioneered by Evans and Lyons (2002, 2005), that tries to explain and empirically forecast exchange rate movements based on a microstructure approach. The microstructure approach assumes that, apart from common knowledge macroeconomic information (macro approach), heterogeneous beliefs are essential for exchange rate determination. In a hybrid view, macroeconomic information influences the exchange rate directly and indirectly through order flow which reveals price-relevant private information such as, for example, heterogeneous interpretations of news or changes in expectations (Rime et al., 2010).[52] Evans and Lyons (2002) provide a theoretical model that integrates both approaches and find empirically that adding order flow as an explanatory variable to a regression of changes in exchange rates on IRDs, serving as a proxy for public macroeconomic information, increases the R-squared from 1%-5% to 40%-60%. As Evans and Lyons (2005) note, order flow data have not only explanatory but also forecasting power for the exchange rate if the market learns gradually from order flow information. Following the out-of-sample studies by Evans and Lyons (2005) and Rime et al. (2010), order flow is a powerful predictor for exchange rate fluctuations. Like order flow information the CME futures position data are not discovered by the market immediately and therefore do not constitute public information. The U.S. Commodity Futures Trading Commission provides the data with a delay of some days (usually three days).

Table 9: Granger Causality Test: Which Variables "Granger-cause" Carry Trade Positions?

Sample / Regime	Variable excluded				
	IRD_{USD} / IRD_{EUR}	VIX	ΔFX_{USD} / ΔFX_{EUR}	Y_{USD} / ΔY_{EUR}	ΔP_{USD} / ΔP_{EUR}
$A^{d=3}_{USD}$					
H	3.062	2.490	20.562***	7.204	1.602
L	3.455	3.565	24.868***	7.102	16.149***
$B^{d=3}_{EUR}$					
H	2.810	0.942	1.762	2.448	0.139
L	17.838***	4.398	0.144	4.067	2.586
$C^{d=3}_{USD}$					
H	4.422	3.373	22.353***	6.040	3.702
L	3.474	1.234	23.797***	5.237	10.955**
$D^{d=3}_{USD}$					
H	4.564	3.910	21.161***	6.956	1.568
L	6.349	1.841	16.781***	2.855	8.912*
$E^{d=3}_{USD}$					
H	7.098	3.167	33.024***	8.922*	2.298
L	4.998	4.376	29.220***	6.690	18.901***

Notes: The samples and variables are described in Table (4). Observations for which the threshold variable lies above the threshold value are assigned to the H-regime; for values below the threshold values, the observations are included in the L-regime. The threshold values are given in Table (7). */**/*** denotes significance of the Chi-squared value at 10%, 5% and 1% level, respectively.

[52] Order flow is defined as the net of buyer and seller initiated currency transactions. Hence, it is a measure of net buying pressure (Evans and Lyons, 2002).

In a second step, the predictive power of all other variables on carry trade positions is determined. The findings are displayed in Table (9). They suggest that exchange rate movements are very important for anticipating carry trade activities, independent of the regime, except for sample $B^{d=3}_{EUR}$. It is therefore more likely that movements in the exchange rate precede position data rather than vice versa. This result is in line with the findings reported by Mogford and Pain (2006).[53] The results indicate a basic form of trend-following behavior among the speculative traders at the CME. Movements in the exchange rate FX_{EUR} do not "Granger-cause" the CTR ratio,[54] but the interest-rate differential and the CTR ratio seem to “Granger-cause” each other in the L-regime (see also Table 8). This might be due to the calculation of the CTR ratio with the IRD as its numerator.

Moreover, in all samples, movements in P_{USD} help to predict position data in periods with $IRD^{d=3}_{USD}$ below the threshold value. The stock market may serve as a proxy for liquidity constraints, determining the value of investor collateral portfolios.

6 Summary and Conclusions

We choose a multivariate time series model to assess (unexpected) movements in carry trade positions. Our model contains variables that determine the profitability of carry traders’ investment. This chapter examines how shocks to these variables affect carry traders’ behavior and vice versa. The set of variables consists of the interest-rate differential (the so-called “carry”), the nominal exchange rate, the VIX index to capture risk sentiment, bond yields to proxy investment returns and the stock market index to model possible liquidity constraints.

Because preliminary analyses of the carry point to a regime-dependent relationship between the variables, a multivariate threshold model is estimated. This specification can account for possible changes in the dynamic behavior of carry trade activities depending on the size of the interest-rate differential.

By analyzing the generalized impulse response functions (GIRFs) of the model, we find the following main results: First, carry trade positions are driven to a large extent by the expected risk on financial markets and the exchange rate. The response of key financial and macroeconomic variables to shocks depends on the size of the carry. These differences then affect carry trade positions. We conclude that in times with a large carry a positive one-standard deviation shock to the carry itself is not enough to compensate investors for the increased risk and that the higher probability of a crash may be due

[53] Klitgaard and Weir (2004) also obtain a statistically significant test statistic for the CHF, but not for most other currencies.

[54] We assume that the CTR ratio is an important indicator for carry traders to adjust their positions. However, as long as investors do not follow strictly this indicator we cannot rule out potential feedback trading.

to fundamentals such as the inflation-rate differential. Moreover, we find that the CHF appreciates instantaneously against the USD in times with high interest-rate differentials (IRDs). This result is in line with the prediction of uncovered interest rate parity (UIP). However, the result does not hold in the regime with low IRDs.

Second, liquidity constraints can also be important, whereas the carry itself plays only a minor role. Third, a sudden unwinding of carry trades has a significant impact on the nominal exchange rate, independent of the size of the IRD. We conclude that carry traders can indeed play a crucial role in determining the nominal exchange rate in the short run and medium run as suggested by Roth (2007).

Finally, we show that the GIRFs of models containing the US dollar/Swiss franc exchange rate are broadly similar to those with the euro/Swiss franc exchange rate, although the proxy for carry trade positions differs.

Furthermore, according to the results of Granger causality tests, past position data help to predict nominal exchange rate fluctuations in periods with low-interest-rate differentials. However, we find that the exchange rate has very high predictive power for carry trade activities when the USD serves as the target currency. From this result we conclude that speculative traders at the CME mainly follow a feedback trading strategy.

Appendix 1.A Additional Tables and Figures

1.A.1 Additional Tables

Table 10: ARCH Test Results with USD as target currency

Dependent Variable	ARCH(1)	ARCH(2)	ARCH(4)
ΔFX_{USD}	0.559	7.213**	11.118**
ΔP_{USD}	42.735***	42.642***	55.007***
VIX	3.015*	5.083*	15.935***
IRD_{USD}	6.181**	10.394***	10.669**
Y_{USD}	4.838**	4.794*	27.997***
Carry Trade Positions			
CTF_{USD}	0.355	1.872	7.328

Notes: The model is estimated with four lags from 1995/03/28 until 2008/06/24. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Table 11: ARCH Test Results with EUR as target currency

Dependent Variable	ARCH(1)	ARCH(2)	ARCH(4)
ΔFX_{EUR}	0.112	17.604***	17.728***
ΔP_{EUR}	41.793***	42.233***	56.998***
VIX	2.997*	4.845*	9.356*
IRD_{EUR}	25.200***	26.300***	32.006***
ΔY_{EUR}	1.434	1.644	5.558
Carry Trade Positions			
CT_{EUR}	14.078***	14.185***	15.454***

Notes: The model is estimated with two lags from 1999/01/06 until 2008/06/25. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Table 12: Granger Causality Test Results for Sample $A^{d=3}_{USD}$

Variable excluded/ Regime	Variable					
	IRD_{USD}	VIX	CTF_{USD}	ΔFX_{USD}	Y_{USD}	ΔP_{USD}
IRD_{USD}						
H		3.062	3.714	4.129	8.496*	6.521
L		3.455	5.418	4.880	6.427	3.539
VIX						
H	1.116		2.490	0.894	1.455	1.843
L	15.669***		3.565	3.862	5.496	0.867
CTF_{USD}						
H	10.560**	9.865**		2.444	3.758	6.493
L	1.390	17.172***		10.206**	14.241***	16.136***
ΔFX_{USD}						
H	4.689	0.244	20.562***		2.136	3.759
L	3.995	3.887	24.868***		6.240	2.798
Y_{USD}						
H	6.440	5.896	7.204	10.900**		9.964**
L	9.368*	10.688**	7.102	5.400		18.378***
ΔP_{USD}						
H	6.396	6.270	1.602	1.994	1.265	
L	11.770**	12.330**	16.149***	13.902***	7.204	
all but own lags						
H	27.068	28.880*	40.000***	23.204	18.022	28.460*
L	42.165***	51.921***	59.574***	35.185**	35.218**	38.106***

Notes: The samples and variables are described in Table (4). The H-regime includes observations where the threshold variable $IRD^{d=3}_{USD}$ is greater or equal to 2.63%. The L-regime includes observations where the threshold variable $IRD^{d=3}_{USD}$ is smaller than 2.63%. */**/*** denotes significance of the Chi-squared value at 10%, 5% and 1% level, respectively.

Table 13: Granger Causality Test Results for Sample $B_{EUR}^{d=3}$

Variable excluded/ Regime	Variable					
	IRD_{EUR}	VIX	CT_{EUR}	ΔFX_{EUR}	ΔY_{EUR}	ΔP_{EUR}
IRD_{EUR}						
H		0.793	2.810	1.522	1.030	1.790
L		3.856	17.838***	1.165	2.529	1.532
VIX						
H	5.886*		0.942	1.878	1.543	1.967
L	1.588		4.398	4.780*	1.395	2.461
CT_{EUR}						
H	1.575	0.257		0.373	7.325**	0.617
L	22.717***	1.862		8.281**	7.832**	11.539***
ΔFX_{EUR}						
H	0.174	1.662	1.762	3.745		5.993**
L	5.850*	2.362	0.144	3.694		0.751
ΔY_{EUR}						
H	0.737	2.329	2.448	4.303		0.018
L	6.916**	1.915	4.067	3.975		0.449
ΔP_{EUR}						
H	3.053	1.701	0.139	1.479	6.243**	
L	0.655	0.153	2.586	0.146	0.187	
all but own lags						
H	13.100	13.992	12.948	7.331	26.838***	11.821
L	45.636***	10.682	29.216***	26.503***	12.457	20.118**

Notes: The samples and variables are described in Table (4). The H-regime includes observations where the threshold variable $IRD_{EUR}^{d=3}$ is greater or equal to 1.84%. The L-regime includes observations where the threshold variable $IRD_{EUR}^{d=3}$ is smaller than 1.84%. */**/*** denotes significance of the Chi-squared value at 10%, 5% and 1% level, respectively.

1.A.2 Additional Figures

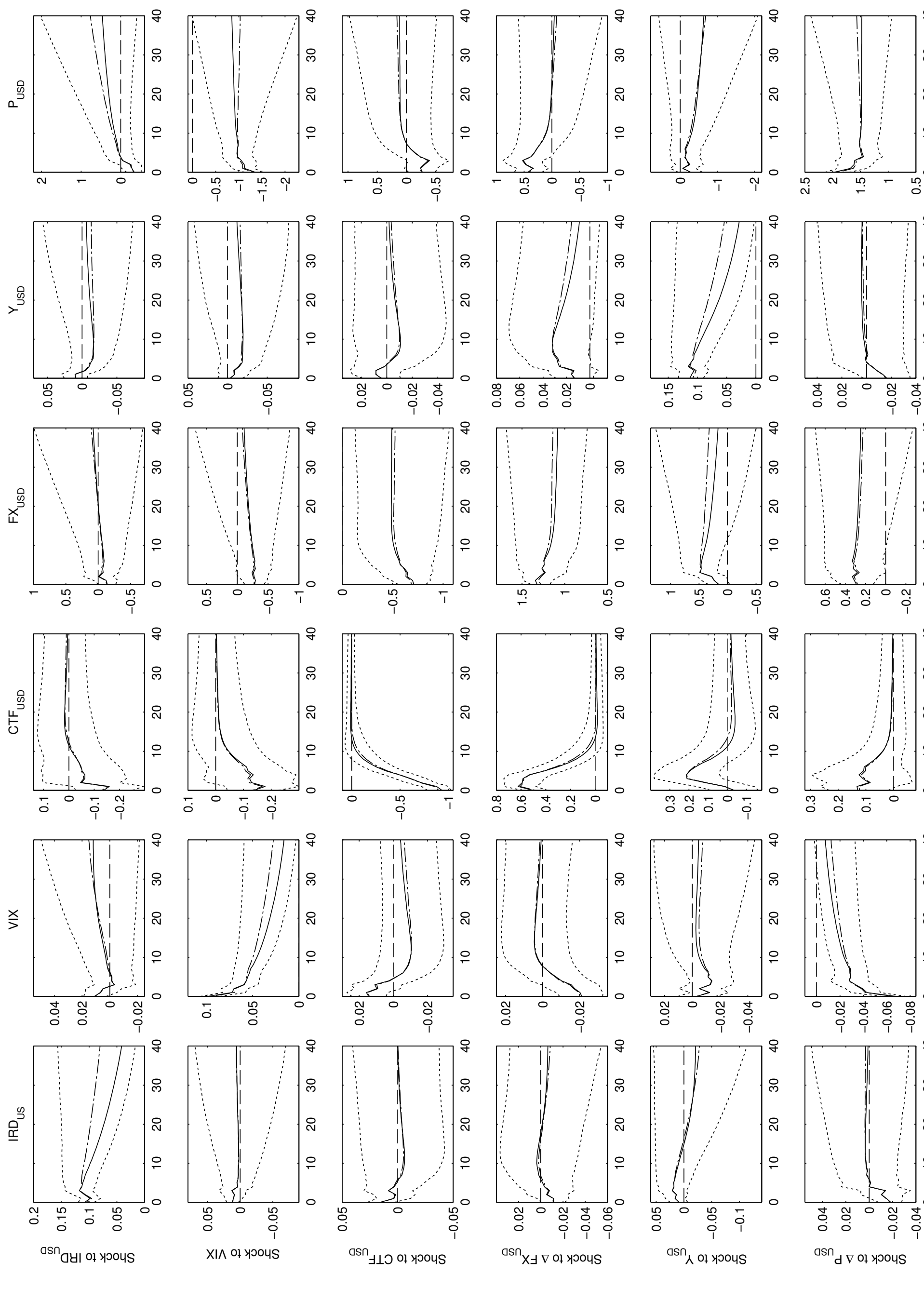


Figure 12: Sample $A^{d=3}_{USD}$: (Accumulated) generalized impulse response functions of the H-regime for all variables. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample A, see Table (4). Number of observations: 418 (H-regime) & 270 (L-regime)

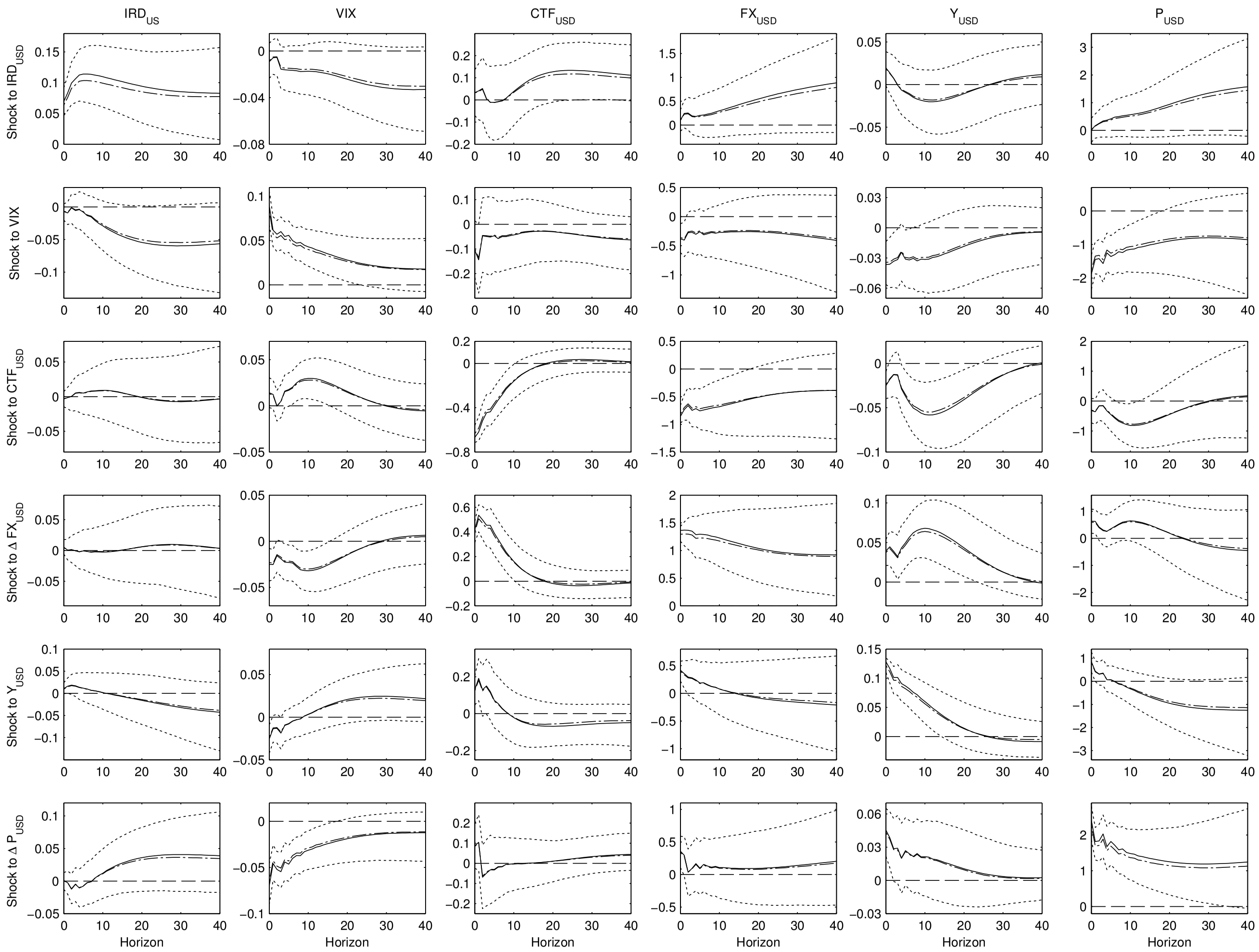


Figure 13: Sample $A_{USD}^{d=3}$: (Accumulated) generalized impulse response functions of the L-regime for all variables. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample A, see Table (4). Number of observations: 418 (H-regime) & 270 (L-regime)

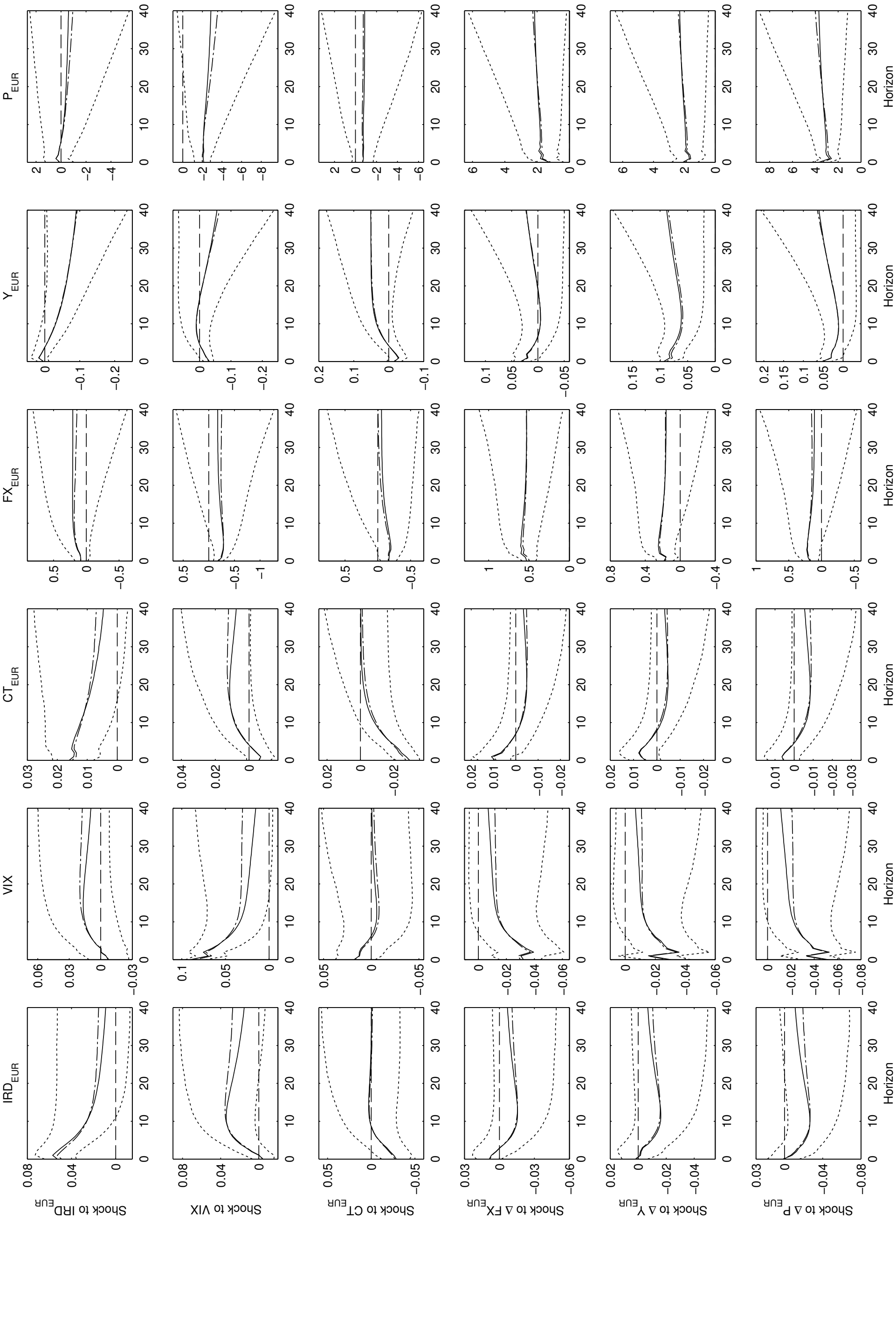


Figure 14: Sample $\mathrm{B}^{d=3}_{EUR}$: (Accumulated) generalized impulse response functions of the H-regime for all variables. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample B, see Table (4). Number of observations: 125 (H-regime) & 367 (L-regime)

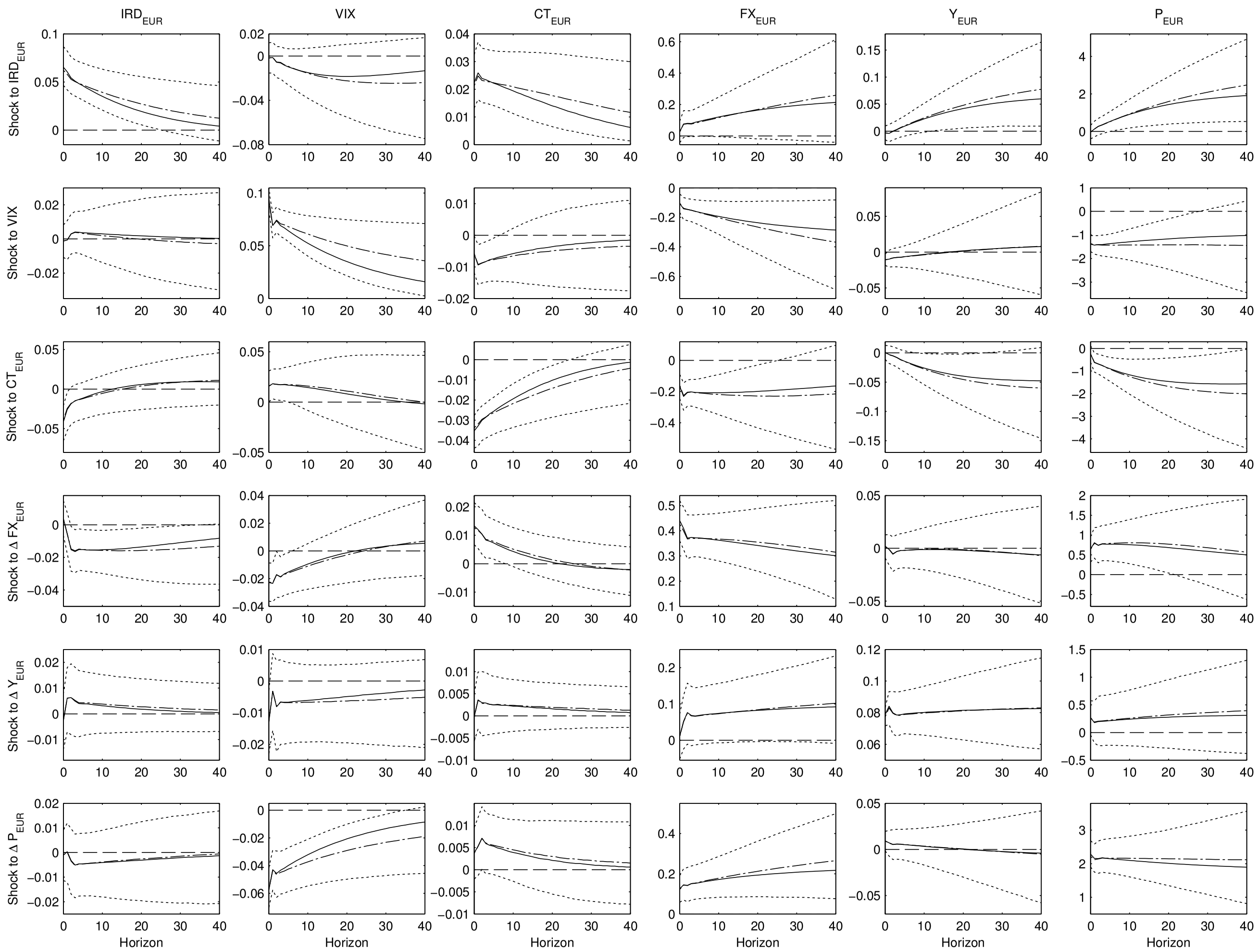


Figure 15: Sample $B_{EUR}^{d=3}$: (Accumulated) generalized impulse response functions of the L-regime for all variables. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample B, see Table (4). Number of observations: 125 (H-regime) & 367 (L-regime)

Bibliography

Bernholz, Peter and Peter Kugler, “Financial market integration in the early modern period in Spain: Results from a threshold error correction model,” *Economics Letters*, 2011, *110* (2), 93–96.

Brunnermeier, Markus K., Stefan Nagel, and Lasse H. Pedersen, “Carry Trades and Currency Crashes,” *NBER Macroeconomics Annual*, 2009, *23* (1), 313–348.

Burnside, Craig, Martin Eichenbaum, Isaac Kleshchelski, and Sergio T. Rebelo, “The Returns to Currency Speculation,” NBER Working Paper 12489, National Bureau of Economic Research, Cambridge, MA August 2006.

Cai, Jun, Yan-Leung Cheung, Raymond S. K. Lee, and Michael Melvin, ““Once-in-a-generation“ yen volatility in 1998: fundamentals, intervention, and order flow,” *Journal of International Money and Finance*, 2001, *20*, 327–347.

Canjels, Eugene, Gauri Prakash-Canjels, and Alan M. Taylor, “Measuring Market Integration: Foreign Exchange Arbitrage and the Gold Standard, 1879-1913,” *The Review of Economics and Statistics*, 2004, *86* (4), 868–882.

Christiano, Lawrence J., Martin Eichenbaum, and Robert Vigfusson, “What Happens after a Technology Shock?,” NBER Working Paper 9819, National Bureau of Economic Research, Cambridge, MA July 2003.

Christiansen, Charlotte, Angelo Ranaldo, and Paul Söderlind, “The time-varying systematic risk of carry trade strategies,” *Journal of Financial and Quantitative Analysis*, 2011, *1* (1), 1–35.

Clarida, Richard, Josh Davis, and Niels Pedersen, “Currency carry trade regimes: beyond the Fama regression,” *Journal of International Money and Finance*, 2009, *28* (8), 1375–1389.

Darvas, Zsolt, “Leveraged carry trade portfolios,” *Journal of Banking & Finance*, 2009, *33*, 944–957.

Dickey, David A. and Wayne A. Fuller, “Distribution of the Estimators for Autoregressive Time Series with a Unit Root,” *Journal of the American Statistical Association*, 1979, *74*, 427–431.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock, “Efficient Tests for an Autoregressive Unit Root,” *Econometrica*, 1996, *64*, 813–836.

Engel, Charles, “The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence,” *Journal of Empirical Finance*, 1996, *3*, 123–192.

Engle, Robert, “GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics,” *Journal of Economic Perspectives*, 2001, *15* (4), 157–168.

Evans, Martin D. D. and Richard K. Lyons, “Order Flow and Exchange Rate Dynamics,” *The Journal of Political Economy*, 2002, *110* (1), 170–180.

_ **and** _ , "Meese-Rogoff Redux: Micro-Based Exchange-Rate Forecasting," *American Economic Review*, 2005, *95* (2), 405–414.

Fama, Eugene F., "Forward and Spot Exchange Rates," *Journal of Monetary Economics*, 1984, *14*, 319–338.

Farhi, Emmanuel, Samuel P. Fraiberger, Xavier Gabaix, Romain Ranciere, and Adrien Verdelhan, "Crash Risk in Currency Markets," NBER Working Paper 15062, National Bureau of Economic Research, Cambridge, MA June 2009.

Gagnon, Joseph E. and Alain P. Chaboud, "What Can the Data Tell Us about Carry Trades in Japanese Yen?," International Finance Discussion Papers 899, Board of Governors of the Federal Reserve System 2007.

Galati, Gabriele, Alexandra Heath, and Patrick McGuire, "Evidence of carry trade activity," *BIS Quarterly Review*, September 2007.

Goncalves, Silvia and Lutz Kilian, "Bootstrapping autoregressions with conditional heteroskedasticity of unknown form," *Journal of Econometrics*, 2004, *123*, 89–120.

Granger, Clive W. J., "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 1969, *37* (3), 424–438.

Hattori, Masazumi and Hyun Song Shin, "The Broad Yen Carry Trade," IMES Discussion Paper Series E-19, Bank of Japan 2007.

Jordà, Oscar and Alan M. Taylor, "The Carry Trade and Fundamentals: Nothing to Fear But FEER Itself," NBER Working Paper 15518, National Bureau of Economic Research, Cambridge, MA November 2009.

Jurek, Jakub W., "Crash-Neutral Currency Carry Trades," AFA 2010 Atlanta Meetings Paper, The American Finance Association 2007.

Kilian, Lutz, "Confidence Intervals for Impulse Responses under Departures from Normality," *Econometric Reviews*, 1998a, *17* (1), 1–29.

_ , "Small-Sample Confidence Intervals for Impulse Response Functions," *The Review of Economics and Statistics*, 1998b, *80* (2), 218–230.

Klitgaard, Thomas and Laura Weir, "Exchange Rate Changes and Net Positions of Speculators in the Futures Market," Economic Policy Review, Federal Reserve Bank of New York 2004.

Kohler, Daniel, "Carry Trades: Betting Against Safe Haven," Discussion Paper 2007-12, University of St. Gallen Department of Economics December 2007.

Koop, Gary, M. Hashem Pesaran, and Simon Potter, "Impulse response analysis in nonlinear multivariate models," *Journal of Econometrics*, 1996, *74* (1), 119–147.

Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt, and Yongcheol Shin, "Testing the Null Hypothesis of Stationary against the Alternative of a Unit Root," *Journal of Econometrics*, 1992, *54*, 159–178.

Lustig, Hanno, Nick Roussanov, and Adrien Verdelhan, "Common Risk Factors in Currency Markets," *Review of Financial Studies*, 2011, *24* (11), 3731–3777.

McGuire, Patrick and Christian Upper, "Detecting FX carry trades," *BIS Quarterly Review*, September 2007.

Meese, Richard A. and Kenneth Rogoff, "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?," *Journal of International Economics*, 1983, *14*, 3–24.

Mogford, Caroline and Darren Pain, "The information content of aggregate data on financial futures positions," Quarterly Bulletin, Bank of England 2006.

Ng, Serena and Pierre Perron, "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica*, 2001, *69* (6), 1519–1554.

Nishigaki, Hideki, "Relationship between the yen carry trade and the related financial variables," *Economics Bulletin*, 2007, *13*, 1–7.

Pesaran, M. Hashem and Yongcheol Shin, "Cointegration and speed of convergence to equilibrium," *Journal of Econometrics*, 1996, *71*, 117–143.

_ **and** _ , "Generalized impulse response analysis in linear multivariate models," *Economics Letters*, 1998, *58*, 17–29.

Phillips, Peter C. B. and Pierre Perron, "Testing for a Unit Root in Time Series Regression," *Biometrika*, 1988, *75*, 335–346.

Plantin, Guillaume and Hyun Song Shin, "Carry Trades and Speculative Dynamics," Unpublished Manuscript, Princeton University and Toulouse School of Economics 2010.

Ranaldo, Angelo and Paul Söderlind, "Safe Haven Currencies," *Review of Finance*, 2010, *14* (3), 385–407.

Rime, Dagfinn, Lucio Sarno, and Elvira Sojli, "Exchange rate forecasting, order flow and macroeconomic information," *Journal of International Economics*, 2010, *80* (1), 72–88.

Roth, Jean-Pierre, "Die Rolle des Franken-Wechselkurses im geldpolitischen Konzept der Schweizerischen Nationalbank," Referat im Rahmen des FEW-HSG, Universität St. Gallen Dezember 2007.

Runkle, David E., "Vector Autoregressions and Reality," *Journal of Business and Economic Statistics*, 1987, *5*, 437–442.

Sims, Christopher A., "Macroeconomics and Reality," *Econometrica*, 1980, *48* (1), 1–48.

_ **, James H. Stock, and Mark W. Watson**, "Inference in Linear Time Series Models with some Unit Roots," *Econometrica*, 1990, *58* (1), 113–144.

Stock, James H. and Mark W. Watson, "Vector Autoregressions," *Journal of Economic Perspectives*, 2001, *15* (4), 101–115.

Tsay, Ruey S., "Testing and Modeling Multivariate Threshold Models," *Journal of the American Statistical Association*, 1998, *93* (443), 1188–1202.

Wagner, Christian, "Risk-Premia, Carry-Trade Dynamics, and Speculative Efficiency of Currency Markets," Working Paper Series 143, Austrian National Bank 2008.

Chapter 2

Commodity Price Shocks and the Business Cycle: Structural Evidence for the U.S.

Abstract

This chapter evaluates the importance of commodity price shocks in the U.S. business cycle. Therefore, we extend the standard set of identified shocks to include unexpected changes in commodity prices. The resulting SVAR shows that commodity price shocks are a very important driving force of macroeconomic fluctuations (second only to investment-specific technology shocks), particularly with respect to inflation. In addition, we find that the systematic contractionary monetary policy feedback rule to sudden increases in commodity prices helped the Fed to achieve price stability in the long run, yet at the cost of a significant economic downturn in output and per-capita hours.

Keywords: Business Cycles, Commodity Price Shocks, Structural VAR

1 Introduction

What are the sources of the U.S. business cycle? In recent years, a great body of research has addressed this question. The set of structural shocks often considered includes (neutral) technology shocks (Galí, 1999), investment-specific technology shocks (Fisher, 2006), and monetary policy shocks (Christiano et al., 1996). Alongside this debate, the macroeconomic effects of unexpected changes in commodity prices — primarily energy prices — have gained a great deal of attention (Kilian, 2008; Hamilton, 2011). However, to our knowledge, no attempt has been made to estimate these four structural shocks in a single VAR model so far. This seems surprising, given that the interactions among the underlying factors are likely to be important (Blanchard and Galí, 2010). In particular, Bernanke et al. (1997) argue that a substantial part of the recessionary effects of oil price shocks in the 1970s was not due to the direct impact of higher producer prices (as suggested by Blinder and Rudd, 2008), but rather due to the systematic contractionary response of the Federal Reserve.[1] In addition, this period was characterized by strong growth in investment-specific technology, while total factor productivity growth was particularly weak (Greenwood et al., 1997). Therefore, we explicitly identify neutral and investment-specific technology shocks when analyzing the interactions between commodity prices and monetary policy. This strategy allows us to disentangle the effects of numerous macroeconomic disturbances that might have played an important role in the U.S. postwar era.

The main aim of this chapter is to quantify the relative importance of the above-mentioned shocks in the U.S. business cycle. Thus, we develop a 9-dimensional VAR with four identified structural shocks. We use standard long-run restrictions (Galí, 1999; Fisher, 2006) in order to identify neutral and investment-specific technology shocks. Monetary policy and commodity price shocks are identified using short-run restrictions, along the lines suggested by Christiano et al. (1996) and Rotemberg and Woodford (1996), respectively. In addition, our estimation strategy explicitly accounts for two types of distortions that may be responsible for the extreme sensitivity of the hours response to (neutral) technology shocks; i.e., the omitted-variable bias (Christiano et al., 2003) and the low-frequency bias (Fernald, 2007; Canova et al., 2010).

The main result of our chapter is that commodity price shocks are a very important driving force of the U.S. business cycle, second only to investment-specific technology shocks. In particular, we find that commodity price shocks explain a large share of cyclical movements in inflation. Sudden variations in the relative price of investment goods are the primary determinant of business cycle fluctuations in output and per-capita hours. Neutral technology shocks and monetary policy shocks, on the other hand, seem less

[1] See also the discussion by Barsky and Kilian (2002) for the 1970s and the survey by Harris et al. (2009) for the 2000s as well as the references therein.

relevant at business cycle frequencies. At low frequencies, however, neutral technology shocks do play an important role in explaining output variability.

We also examine the impulse response functions triggered by each of the four identified structural shocks. A sudden increase in *commodity prices* is characterized by significant U-shaped responses in output, consumption and per-capita hours. Most notably, the inflation rate displays a significant spike, followed by a rapid return to the initial level. The sudden surge in the inflation rate prompts the Fed to elevate the nominal interest rate. Results of a counterfactual exercise (in the style of Bernanke et al., 1997) indicate that the systematic contractionary response helped the Federal Reserve to achieve price stability in the long run, yet at the cost of a significant economic downturn in output and per-capita hours. The estimated dynamics turn out to be robust when we control for movements in external demand. Furthermore, we find that the estimated impulse response functions to *neutral technology* shocks, *investment-specific technology* shocks, and *monetary policy* shocks show only minor differences compared to those obtained by Altig et al. (2011) or Ravn and Simonelli (2008). In particular, the response of per-capita hours to neutral technology shocks is positive and marginally significant. This result is surprising, given that we control for low-frequency movements in the data (Canova et al., 2010). Further investigations reveal that this result is very robust to the treatment of the data as long as the size of the information set is sufficiently large. On the other hand, if the information set is small, the impact response of per-capita hours is indeed extremely sensitive to the treatment of the data. This result, which is in line with the evidence found by Forni and Gambetti (2011), confirms our choice to estimate a large-scale SVAR.

The sub-sample properties of our model are consistent with Blanchard and Galí (2010). We find that the effects of a commodity price shock on output and the inflation rate are milder in the post-Volcker period, but the impulse responses remain statistically significant at the 10% level. Several further robustness checks confirm the findings of our model. In particular, we examine robustness to the choice of the lag length, the identification of the commodity price shock, the specific commodity price index used, and the inclusion of an external demand shock (Abbritti and Weber, 2010).

The remainder of this chapter is organized as follows. Section 2 presents the identification and estimation strategies. Section 3 presents the results. Section 4 performs several robustness checks. Section 5 concludes.

2 Identification and Estimation Strategy

We estimate a VAR with four identified structural shocks (neutral technology, investment-specific technology, monetary policy, and commodity prices). Our strategy adopts standard identifying assumptions. We modify the code by Altig et al. (2011) to estimate the coefficients and compute the confidence intervals with a non-parametric bootstrap.[2] There are two novel aspects in our analysis. First, we add commodity price shocks to the standard set of identified structural shocks. Second, our estimation strategy explicitly accounts for two types of distortions that may be responsible for the extreme sensitivity of the hours response to (neutral) technology shocks. Therefore, we estimate the resulting SVAR using a large information set. In particular, the inclusion of the consumption and the investment share has been proven crucial to minimize omitted-variable bias toward a *negative* impact response (Christiano et al., 2003). Furthermore, we apply a one-sided bandpass filter prior to estimation (Canova et al., 2010). The novel feature of our analysis is that we filter not only per-capita hours, but *all* time series that enter the SVAR. This procedure allows us (*i*) to control for low-frequency movements in the data and (*ii*) to maintain spectral coherence (Granger, 1969). As demonstrated by Fernald (2007), these low-frequency movements are likely to distort the estimation toward a *positive* impact response.

2.1 Data

The sample period of this chapter covers aggregate U.S. data between 1955Q3 and 2007Q4.[3] The following variables enter the SVAR: growth in the relative price of investment goods, Δq_t, growth in labor productivity, Δa_t (measured by the ratio of real output to hours per capita in the business sector), the CPI inflation rate, π_t, hours per capita, h_t, the consumption share in output, c_t, the investment share in output, i_t, the employment rate, n_t, the Federal Funds rate, r_t, and the commodity price index "PPI: crude materials for further processing", p_t. We prefer to use this particular commodity price index by the BLS (2012, p. 8) as it appropriately captures the time-varying importance of different raw materials, based on input-output studies by the BEA.[4] All time series are seasonally adjusted (where applicable). Precise definitions can be found in Appendix 2.A (Tables 5 and 6).

The discussion sparked by Galí (1999) and Francis and Ramey (2005) has shown that the response of hours worked to (neutral) technology shocks is extremely sensitive to the treatment of the data. When hours worked enter in first differences, the SVAR typically generates a *negative* impact response. The opposite holds true when the series enters in levels (Christiano et al., 2003, 2004; Uhlig, 2004).[5] Since hours worked are

[2] We thank Lawrence Christiano for making the code available on his website.

[3] The endpoint of our sample marks the start of the Great Recession when the Federal Reserve adopted several unconventional monetary policy measures, which are unlikely to be appropriately captured by our identification procedure.

[4] In contrast, the Thomson Reuters (2010) Continuous Commodity Index, which was calculated backwards until 1956Q4, continuously rebalances the different commodity categories to maintain an equal and time-invariant weight (see also Section 4.5 and Figure 18). The Thomson Reuters/Jefferies (2011) CRB Index, which uses time-varying weights, was calculated backwards only until 1994Q1.

[5] Dedola and Neri (2007) question the above-mentioned sensitivity. Using a sign restriction approach,

borderline stationary, both choices can be justified on the basis of standard unit root tests. Even though the hours series is bounded, the presence of low-frequency movements may prevent the rejection of the null hypothesis of stationarity. These low-frequency movements, sometimes referred to as "long cycles", may be attributed to sectoral changes involving government and non-profit employment or the movement of the baby boom generation through the labor market (Francis and Ramey, 2009).

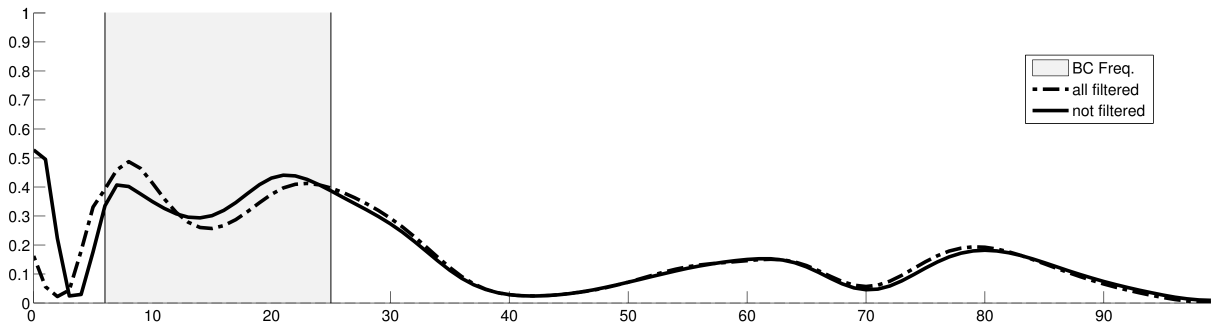
As convincingly demonstrated by Fernald (2007), the presence of "long cycles" in per-capita hours may lead to significant distortions. The author illustrates the low-frequency bias by performing the following counterfactual exercise. Fernald removes the high and medium frequencies from the hours series, reverses their sign, and then adds both series together. Surprisingly, the impact response of hours worked to neutral technology shocks remains positive, although all high and medium frequencies are reversed. This indicates that the positive impact response in a bivariate VAR model is solely driven by the presence of low-frequency movements. Differencing removes the low-frequency movements from the data. This explains why we observe a negative impact response when hours worked are assumed to be non-stationary. Yet, differencing a bounded series (like per-capita hours) may involve misspecification issues (Hamilton, 1994, p. 652). For this reason, Canova et al. (2010) evaluate several alternative filtering devices (e.g., bandpass filter, dummies) that are able to capture these long cycles in the data, but do not induce overdifferencing. Their three-dimensional SVAR identifies a neutral and an investment-specific technology shock. They conclude that all tested filtering methods produce results consistent with Galí (1999) and Francis and Ramey (2005). On impact, per-capita hours fall significantly in response to neutral technology shocks. After this, the function converges monotonically to its initial level.

With this in mind we treat the data as follows. First, we take the natural logarithm of all variables — except for the (net) Federal Funds rate. Next, we difference labor productivity and the relative price of investment goods.[6] Then, in order to control for low-frequency movements in per-capita hours, we employ a one-sided bandpass filter (Christiano and Fitzgerald, 2003) prior to estimation.[7] We prefer this particular filter since agents know only the past (Lucas, 1980). Moreover, we apply the one-sided bandpass filter not only to per-capita hours, but to *all* series considered. Figure (1) illustrates that this procedure allows us to maintain spectral coherence between labor productivity

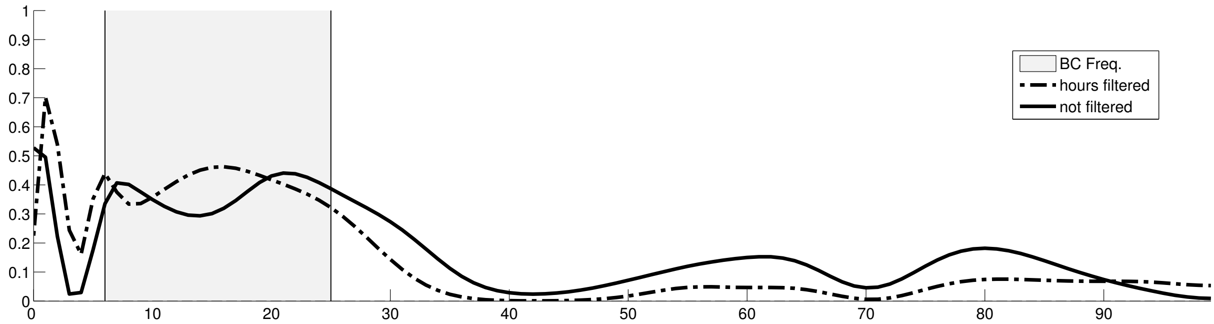
they show that a positive technology shock raises hours worked irrespective of the data transformation.

[6]Our long-run identification strategy of the two technology shocks (see Section 2.2) requires that labor productivity and the relative price of investment goods enter the SVAR in first differences. Hence, we difference these variables first and then apply the bandpass filter. Section 4.1 shows that this procedure is not driving our results.

[7]To be precise, we first remove the drift of the series and then apply the one-sided bandpass filter with following options: $p_l = 2$, $p_u = 52$, $root = 1$, $drift = 0$, $ifilt = 0$, $nfix = -1$, $thet = 1$. Where available, we use data from 1948Q1 to 2007Q4 and then drop the filtered data points prior to 1955Q3. The Federal Funds rate is only available from 1954Q3. Choosing $root = 0$ instead leads to less favorable coherence properties, but leaves the main conclusions unchanged.



(a) all series are filtered



(b) only per-capita hours are filtered

Figure 1: Coherence Analysis – The figure illustrates the coherence between labor productivity growth and per-capita hours, estimated with five lags.

growth and per-capita hours in our benchmark specification. When all series are filtered — as in the top panel — we are able to break the low-frequency comovement and minimize distortions at higher (particularly, at business cycle) frequencies. When only per-capita hours are filtered — as in the bottom panel — we are less successful in breaking the low-frequency comovement and distort the relationship at business cycle frequencies.

Using a trivariate SVAR, we are able to replicate the results of Canova et al. (2010) — both in terms of business cycle variance decomposition and estimated impulse response functions. However, we find significant correlation coefficients between the estimated neutral technology shock and the filtered series of the inflation rate, the Federal Funds rate and the index of commodity prices, respectively, at various leads and lags. Besides, the estimated investment-specific shock is significantly correlated with the filtered series of the inflation rate, the consumption share, and the investment share (see Table 1). Therefore, we extend the trivariate VAR to these additional variables. Thus, we minimize the possibility of omitted-variable bias (Christiano et al., 2003). Furthermore, we note that the filtered series of the employment rate is not significantly correlated with either the neutral or the investment-specific technology shock. Nevertheless, we include this variable as a labor market indicator in our information set. In contrast to, for example, the unemployment rate, the employment rate corresponds to the business sector and, thus, allows us to examine employment adjustment consistently along the extensive vs.

Table 1: Cross Correlations with Technology Shocks

		0	1	2	3	4	5
c_t	lag	0.091	0.080	0.090	0.082	0.094	0.059
	lead	0.091	0.108	0.108	0.069	0.028	0.033
n_t	lag	-0.092	-0.033	-0.018	-0.019	-0.012	-0.001
	lead	-0.092	-0.105	-0.119	-0.085	-0.064	-0.077
r_t	lag	-0.104	-0.052	-0.016	0.021	0.020	0.068
	lead	-0.104	-0.152^{*}	-0.190^{**}	-0.210^{**}	-0.197^{**}	-0.145^{*}
π_t	lag	-0.280^{**}	-0.040	-0.047	-0.122	-0.074	-0.010
	lead	-0.280^{**}	-0.171^{*}	-0.136	-0.136	-0.107	-0.112
i_t	lag	0.001	-0.026	-0.012	-0.033	-0.045	-0.033
	lead	0.001	0.012	0.038	0.076	0.100	0.094
p_t	lag	-0.188^{**}	-0.065	-0.050	-0.062	-0.025	0.003
	lead	-0.188^{**}	-0.154^{*}	-0.134	-0.129	-0.103	-0.100

(a) neutral technology

		0	1	2	3	4	5
c_t	lag	-0.038	-0.143^{*}	-0.118	-0.148^{*}	-0.083	-0.099
	lead	-0.038	-0.017	0.018	0.033	0.044	0.064
n_t	lag	-0.019	0.011	0.022	0.007	0.004	0.003
	lead	-0.019	-0.036	-0.081	-0.111	-0.115	-0.091
r_t	lag	0.053	0.096	0.095	0.021	0.010	-0.018
	lead	0.053	0.003	-0.056	-0.075	-0.048	-0.074
π_t	lag	-0.092	0.188^{**}	0.175^{*}	0.016	0.156^{*}	0.123
	lead	-0.092	-0.012	-0.015	-0.137	-0.100	0.014
i_t	lag	-0.081	0.024	0.031	0.028	0.000	0.016
	lead	-0.081	-0.104	-0.171^{*}	-0.177^{*}	-0.133	-0.103
p_t	lag	-0.007	0.066	0.103	0.086	0.081	0.064
	lead	-0.007	0.000	-0.021	-0.051	-0.002	0.038

(b) investment-specific technology

Notes: The table displays cross correlation coefficients with the two identified technology shocks at leads and lags (± 5 quarters). Stars ($*$, $**$) indicate significance at the 5% and 1% level, respectively.

intensive margin. None of our results are affected if we exclude the employment rate from our system.[8] The extension of the information set is also supported by the outcome of bivariate Granger (1969) causality tests (see Table 2).

2.2 Identification

We estimate four structural shocks using standard identifying assumptions. To begin with, we consider two kinds of technology shocks. Neutral technology shocks (Kydland and Prescott, 1982) are able to replicate the cyclical comovement of output, hours worked, and consumption easily. For this reason, this type of disturbance has played a dominant role in the early RBC literature (King and Rebelo, 1999). More recently, however, investment-specific technology shocks have gained a great deal of attention. This strand of the

[8] Alternatively, we have tested whether we can exclude single time series from our 9-dimensional SVAR. The resulting dynamic correlation pattern is consistent with the one presented in Table (1).

Table 2: Granger Causality Tests

	c_t	n_t	r_t	π_t	i_t	p_t
q_t	0.04	0.73	0.17	0.00	0.64	0.61
a_t	0.81	0.86	0.08	0.07	0.69	0.34
h_t	0.83	0.29	0.00	0.11	0.00	$0.14^{\dagger}$

Notes: The table displays the Granger causality/block exogeneity Wald test statistics when 5 lags are included. We obtain the p-values from bivariate VARs where the residuals of the trivariate model (in rows) are tested against potentially omitted variables (in columns). (†) The null hypothesis between commodity prices and the hours residual can only be rejected when the number of lags is set to 1 and 2 (at the 5% significance level, respectively), or 3 (at the 10% significance level).

literature argues that movements in the relative price of investment goods are not only important in explaining postwar U.S. growth, but also for macroeconomic fluctuations at business cycle frequencies (Greenwood et al., 2000). In addition to these two disturbances, we identify two non-technology shocks — innovations to monetary policy (Christiano et al., 1996) and unexpected changes in commodity prices (Rotemberg and Woodford, 1996). The resulting identification procedure is equivalent to the model of Ravn and Simonelli (2008), with commodity prices instead of government spending.[9] The remaining shocks in the nine-dimensional SVAR model are identified via a recursive ordering scheme.

Consequently, the reduced-form VAR is given by:

$$x_t = a + B(L)x_{t-1} + e_t \tag{1}$$
$$x_t = \left[\begin{array}{ccccc} \Delta q_t & \Delta a_t & z_t & r_t & p_t \end{array} \right]'$$
$$z_t = \left[\begin{array}{ccccc} \pi_t & h_t & c_t & i_t & n_t \end{array} \right]'$$

where $B(L)$ is a lag polynomial of order M. By premultiplication with β_0, one obtains the structural VAR:

$$\beta_0 x_t = \alpha + \beta(L)x_{t-1} + \epsilon_t \tag{2}$$

where ϵ_t denotes the vector of fundamental shocks. The orthogonality assumption implies that its covariance matrix $V_\epsilon = E(\epsilon_t' \epsilon_t)$ is diagonal. Moreover, we normalize the diagonal of β_0 to a 9×1 vector of ones.

Both technology shocks are identified using long-run restrictions (Shapiro and Watson, 1988; Blanchard and Quah, 1989). Following Fisher (2006), we assume that *only* investment-specific technology shocks affect the relative price of investment goods in the long run. The long-run level of aggregate productivity may be affected by both investment-specific *and* neutral technology shocks. No other shock has any long-run effect on the relative price of investment goods or the level of labor productivity (Galí, 1999).

[9]Therefore, we have also estimated our SVAR with real government spending per capita instead of commodity prices. Such a government spending shock, however, does not seem to cause significant movements in U.S. macroeconomic aggregates.

Our identification strategy of the two remaining shocks is based on short-run restrictions. We impose the constraint that no other variable may respond contemporaneously when the Fed's monetary policy — given by the Federal Funds rate — deviates from its linear rule. This presumes that, when setting the nominal interest rate, the Fed's information set includes the contemporaneous values of all other variables included in the SVAR (Christiano et al., 1996). Moreover, we identify the commodity price shock by assuming that all other shocks have no impact on the contemporaneous value of the commodity price index.[10] This assumption was originally developed by Rotemberg and Woodford (1996) in the context of nominal oil price shocks.[11] It is based on the perception that the sources of short-run oil price fluctuations, such as political strife in the Middle East (Kilian, 2008), are exogenous to the U.S. economy. However, this may be incorrect because the U.S. is a large economy, or because economic developments in the U.S. are correlated with global economic activity (Blanchard and Galí, 2010). Indeed, Kilian (2008) argues that oil prices should be treated as endogenous. Nevertheless, he concludes that the *contemporaneous* exogeneity assumption provides a good approximation when working with quarterly data.

When applying this approach to commodity prices, we take into account the possibility that the broad commodity price index may behave differently from the nominal oil price (Alquist et al., 2011). Therefore, Section 4.4 relaxes the contemporaneous exogeneity assumption to allow for immediate responses in the commodity price index to unexpected changes in two main indicators of the U.S. economy (labor productivity growth and per-capita hours). In line with Blanchard and Galí (2010), we find that the estimated results are robust to this alternative identification scheme. In addition, Section 4.3 distinguishes between supply and demand-driven innovations.[12]

Consequently, the process for the Federal Funds rate depends on the current and past values of all other variables, but no other process depends on its current realizations. This implies that the second-last column of the contemporaneous coefficient matrix β_0 consists of zeros, apart from the second-last element which is normalized to unity. The process for the commodity price, on the other hand, depends on the lagged values of commodity prices and all other variables, but not on the current realizations of any other variable. Hence, the last row of β_0 consists of zeros, apart from the last element which is normalized to unity. Furthermore, the order of the variables included in the vector z_t imposes a number of additional short-run restrictions on β_0.

[10]Our identification procedure presumes that commodity price increases and decreases have symmetric effects. This assumption is based on the results of Kilian and Vigfusson (2011) who find no asymmetry in the responses to energy price shocks.

[11]The shape of the impulse responses remains unchanged when we use the West Texas Intermediate spot oil price instead. Quantitatively, however, the broad commodity price index turns out to be more important for the cyclical behavior of U.S. macroeconomic aggregates.

[12]However, as commodity prices are denominated in U.S. dollars, we are not able to identify endogenous (real) exchange rate driven commodity price movements.

2.3 Estimation

The first equation of the structural VAR (Equation 2):

$$p_t = \alpha^p + \sum_{j=1}^{M} \beta^p_{x,j} x_{t-j} + \epsilon^p_t \tag{3}$$

identifies the commodity price shock ϵ^p_t. We estimate Equation (3) using ordinary least squares. The second equation of the SVAR:

$$\begin{aligned}\Delta q_t = \alpha^q &+ \sum_{j=1}^{M} \beta^q_{q,j} \Delta q_{t-j} + \sum_{j=0}^{M-1} \beta^q_{a,j} \Delta^2 a_{t-j} \\ &+ \sum_{j=0}^{M-1} \beta^q_{z,j} \Delta z_{t-j} + \sum_{j=1}^{M-1} \beta^q_{r,j} \Delta r_{t-j} + \sum_{j=0}^{M-1} \beta^q_{p,j} \Delta p_{t-j} + \epsilon^q_t\end{aligned} \tag{4}$$

identifies the investment-specific technology shock ϵ^q_t. The long-run restriction is imposed by differencing all the regressors in x_t apart from the relative investment goods price itself (note that Δ^2 is the second difference operator). Moreover, we exclude the contemporaneous value of the Federal Funds rate from this regression. This implements the short-run assumption on the Fed's information set. Since ϵ^q_t may be correlated with Δa_t (via Equation 5) and z_t (via Equation 7), we estimate Equation (4) with 2SLS. The instruments are a constant, the vector $[\Delta q_{t-j}, \Delta a_{t-j}, z_{t-j}, r_{t-j}, p_{t-j}]_{j=1}^{M}$ and $\hat{\epsilon}^p_t$ (the estimate of ϵ^p_t). The third equation of the SVAR:

$$\begin{aligned}\Delta a_t = \alpha^a &+ \sum_{j=0}^{M} \beta^a_{q,j} \Delta q_{t-j} + \sum_{j=1}^{M} \beta^a_{a,j} \Delta a_{t-j} \\ &+ \sum_{j=0}^{M-1} \beta^a_{z,j} \Delta z_{t-j} + \sum_{j=1}^{M-1} \beta^a_{r,j} \Delta r_{t-j} + \sum_{j=0}^{M-1} \beta^a_{p,j} \Delta p_{t-j} + \epsilon^a_t\end{aligned} \tag{5}$$

identifies the neutral technology shock ϵ^a_t. Note that we difference all regressors — except for Δq_t and Δa_t — and exclude the contemporaneous value of the Federal Funds rate. We estimate Equation (5) using 2SLS, given that ϵ^a_t may depend on z_t (via Equation 7) and q_t (via Equation 4). The instruments employed above are extended to include the estimate of ϵ^a_t; i.e., $\hat{\epsilon}^a_t$. The fourth equation of the SVAR:

$$r_t = \alpha^r - \beta^r_{q,0} \Delta q_t - \beta^r_{a,0} \Delta a_t - \beta^r_{z,0} z_t - \beta^r_{p,0} p_t + \sum_{j=1}^{M} \beta^r_{x,j} x_{t-j} + \epsilon^r_t \tag{6}$$

identifies the monetary policy shock ϵ^r_t. This equation is estimated with ordinary least squares.

Following Altig et al. (2011), we estimate the remaining parameters for the vector z_t. The components of z_t are denoted by z^i_t, $i = 1, \ldots, 5$. The parameters of the first equation are obtained by estimating:

$$z_t^1 = \alpha^1 + \sum_{j=0}^{M} \beta_{q,j}^1 \Delta q_{t-j} + \sum_{j=0}^{M} \beta_{a,j}^1 \Delta a_{t-j} + \sum_{j=1}^{M} \beta_{z,j}^1 z_{t-j} + \sum_{j=1}^{M} \beta_{r,j}^1 r_{t-j} + \sum_{j=0}^{M} \beta_{p,j}^1 p_{t-j} + \epsilon_t^1, \tag{7}$$

and employing the above-used instruments including the vector of estimated shocks [$\hat{\epsilon}_t^p$, $\hat{\epsilon}_t^q$, $\hat{\epsilon}_t^a$]. The second equation extends the set of regressors with z_t^1 and the list of instruments with $\hat{\epsilon}_t^1$. We continue this procedure recursively for all the variables included in z_t.

3 Results

We apply standard lag selection tests to determine the optimal VAR order (M). These tests, however, yield inconsistent results. The information criteria by Akaike (3), Hannan-Quinn (1), and Schwarz (1) indicate a short lag length. Sequential likelihood ratio test, on the other hand, suggest that the VAR order is somewhat larger.[13] Given that our identification strategy of technology shocks is based on long-run restrictions, we set the lag length equal to $M = 5$. Thus, we minimize the possibility of "truncation bias".[14] Nevertheless, our results indicate that the qualitative shape of the impulse responses is robust to the number of lags included. Quantitatively, we find that the business cycle variance of output, per-capita hours, and the inflation rate which is explained by the four identified shocks rises when we increment the number of lags from three to five (see Figure 2). In addition, we test the null hypothesis of zero serial correlation using bootstrapped multivariate Portmanteau (Q) statistics (Altig et al., 2011). On the basis of this test, we do not reject the null hypothesis when the VAR order is set to $M = 5$.

3.1 Dynamic Responses to Structural Shocks

We examine the impulse responses at horizons up to 32 quarters. The graphs depict the responses based on bootstrap sampling over 3,000 replications, where the first 1,000 draws are used to correct for small sample bias and departures from non-normality (Kilian, 1998a,b).[15] The solid line is the median estimate. The gray shaded areas represent the associated 60%, 70%, 80% and 90% non-centered confidence intervals. For the reader's

[13]The standard sequential likelihood ratio test (see, e.g., Lütkepohl, 2005, Chapter 4.2.2) rejects $M = 4$ at the 1% significance level, the modified sequential likelihood ratio test (Sims, 1980) rejects $M = 3$ at the 5% significance level.

[14]Erceg et al. (2005) provide evidence that such finite-ordered VARs are able to approximate the true data generating process sufficiently well. Our choice to use a lag length beyond one year is also supported by the existing literature on energy price shocks (Hamilton and Herrera, 2004).

[15]The Jarque-Bera test statistics reject the null hypothesis that the commodity price shocks and the monetary policy shocks are normally distributed at the 1% significance level.

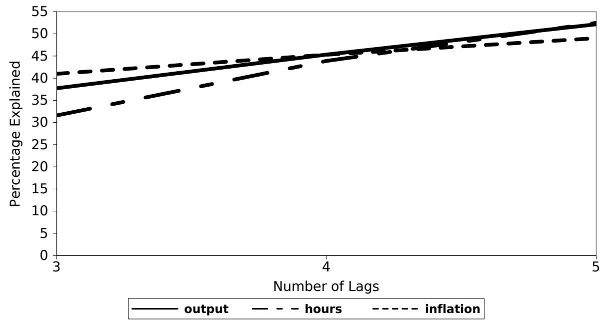


Figure 2: Lag Length and Business Cycle Variance Decomposition – The figure illustrates the share of business cycle variance explained by the four identified structural shocks when the lag length increases from three to five.

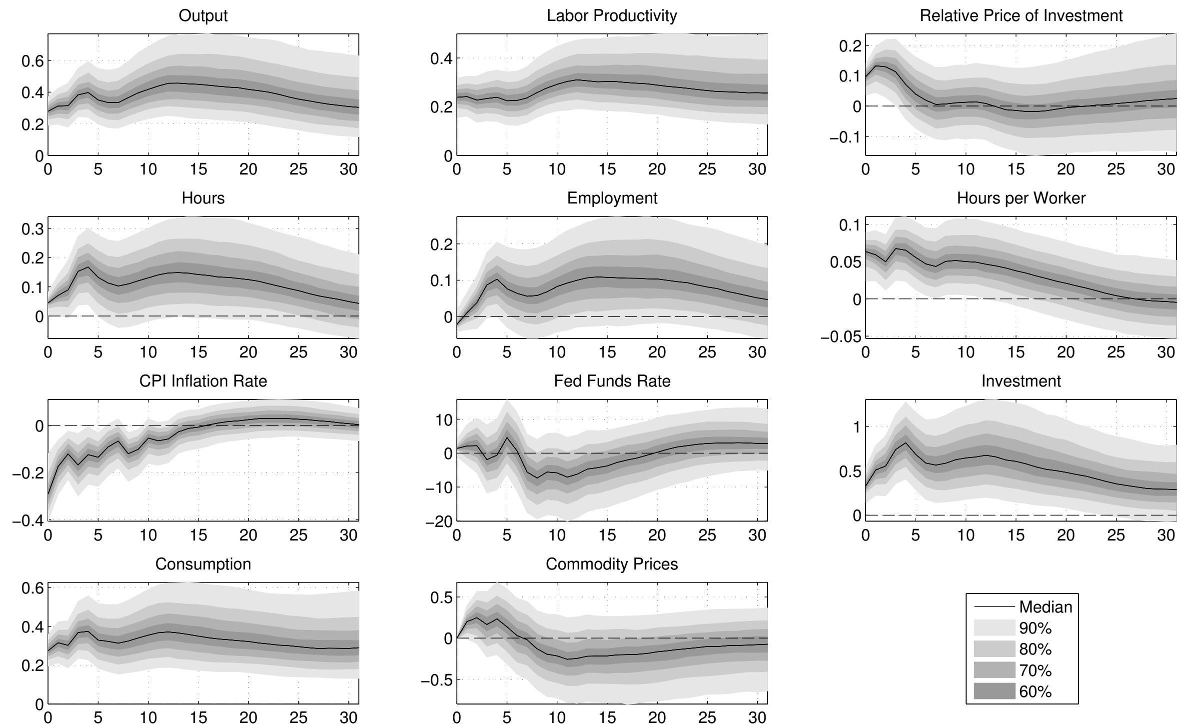
convenience, Figures (3), (5)-(7) contrast the impulse responses of our benchmark specification (a panels) with the impulse responses of the level specification (b panels).

3.1.1 Neutral Technology Shocks

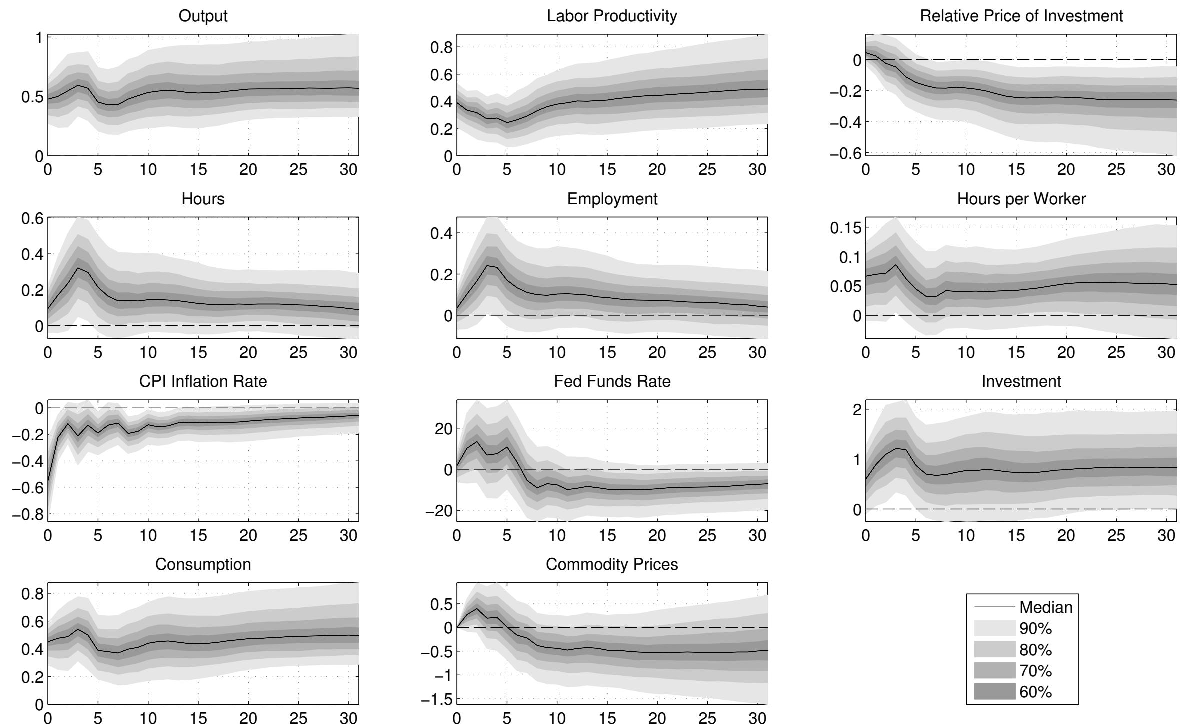
Figure (3) illustrates the impulse response functions to the identified neutral technology shock. We observe that a permanent improvement in labor productivity induces a long-lasting rise in output and consumption. On impact, both variables jump up and then remain well above their original value for the entire time horizon. Moreover, the shock produces a large and protracted hump-shaped response in investment. The inflation rate falls on impact and then asymptotes to its steady-state level within four years. There is also a modest increase in the relative price of investment goods, but the effect disappears relatively quickly. The impulse response of per-capita hours is positive and marginally significant at the 10% level. A very similar response can be observed for the employment rate. Hours per worker, on the other hand, rise on impact and then slowly return to their steady state. Quantitatively, however, the impact of the intensive margin is small. The estimated impulse responses differ only in one important respect from those obtained by Altig et al. (2011). We find that the increase in consumption is not gradual, but rather abrupt. The major implication of our result is that models assuming habit formation in aggregate consumption (Abel, 1990) may not be able to replicate the dynamics of the U.S. economy.

Given that the response of hours worked to neutral technology shocks is extremely sensitive to the treatment of the data, we also examine the impact of alternative specifications in the following. Figure (4) displays the hours response in our benchmark specification, the corresponding level specification, the corresponding difference specification, a dummy specification, and in the level specification using the corresponding Francis and Ramey (2009) hours time series.[16] When the information set is large — as in the left panel — we

[16] *Level* specification: All series enter in levels, but labor productivity and the relative price of investment goods enter in first differences. *Difference* specification: Also per-capita hours enter in first



(a) Benchmark Specification



(b) Level Specification

Figure 3: The figure illustrates the impulse responses to a neutral technology shock.

differences. *Dummy* specification: We extend the level specification to include a time trend and two structural breaks in level and trend at the dates 1973Q2 and 1997Q2 (see Fernald, 2007). The Francis and Ramey (2009) hours time series is taken from Valerie A. Ramey's website, which is gratefully acknowledged.

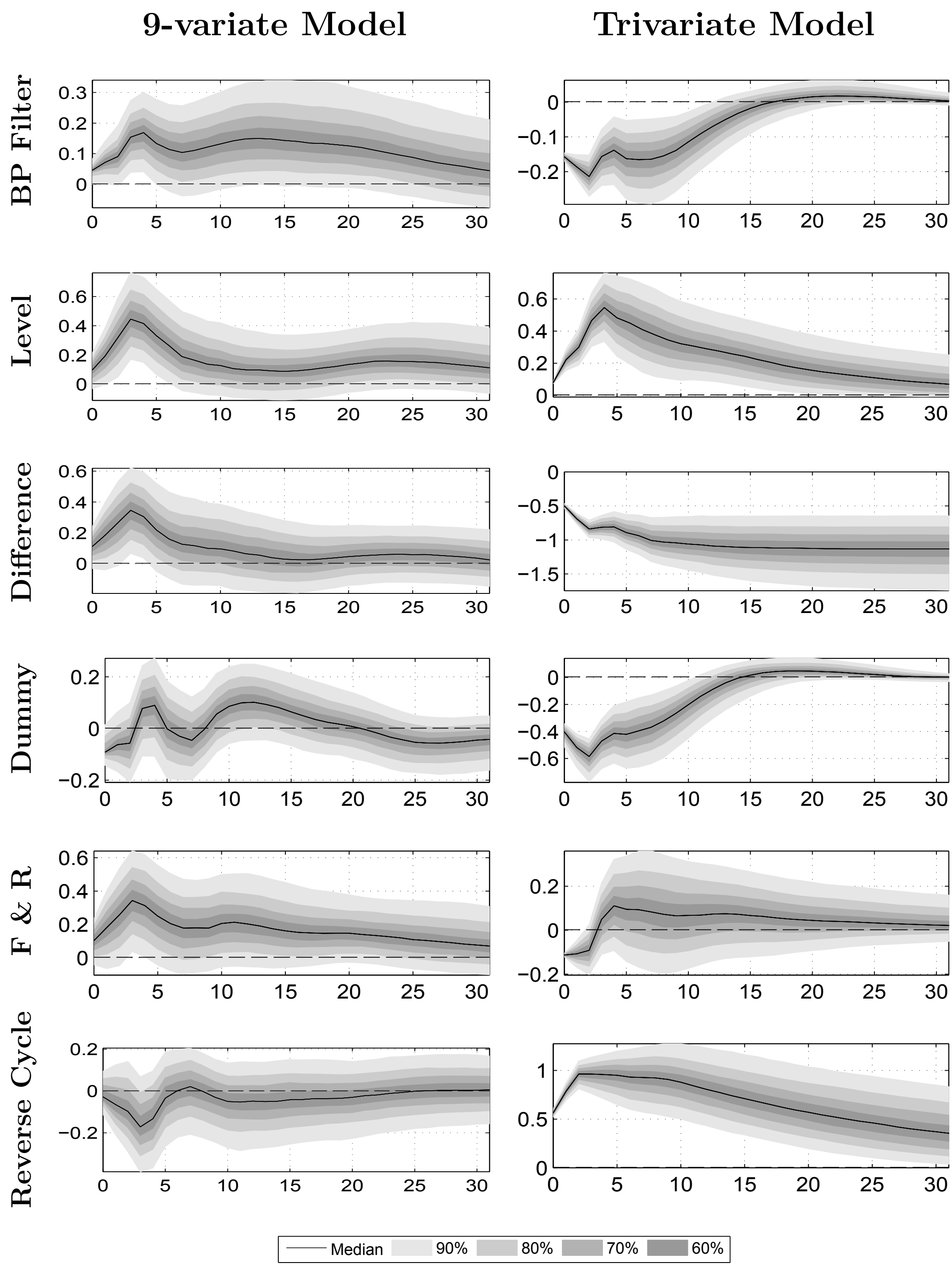


Figure 4: The figure illustrates the low-frequency bias by the means of the per-capita hours response to a neutral technology shock.

observe that our results are very robust. Across all specifications, the dynamic response is positive and marginally significant at the 10% level. Only the dummy specification predicts a negative response during the first few quarters, but the confidence intervals are wide. Interestingly, the hours response flips horizontally when all high and medium frequencies in per-capita hours are reversed (Fernald, 2007).[17] These results indicate that the low-frequency bias (present in the level specification) and the misspecification error (induced by overdifferencing) become less important when the information set is sufficiently large. This conclusion is consistent with the results of Christiano et al. (2003) and Forni and Gambetti (2011).[18]

On the other hand, the right panel of Figure (4) shows the impulse responses when the information set is reduced to three variables $\{q_t, a_t, h_t\}$. We observe that, in this case, the hours response becomes indeed extremely sensitive to the treatment of the data. If we remove the low-frequency movements, either by applying the one-sided bandpass filter, by taking first differences, by including a time trend and two structural breaks in level and trend, or by including the corresponding Francis and Ramey (2009) hours time series in the level specification, the hours response is significantly *negative*. This confirms the results of Galí (1999) and Canova et al. (2010). Instead, the response is significantly *positive* if per-capita hours enter the VAR in levels, thus echoing the findings of Christiano et al. (2003, 2004). Moreover, we are able to replicate the counterfactual exercise conducted by Fernald (2007). Even if all high and medium frequencies in per-capita hours are reversed, the trivariate SVAR generates a significant positive response. In summary, these results imply that the (downward) omitted-variable bias and the (upward) low-frequency bias lead to significant distortions only when the information set is insufficiently small.

The reason why the results of Galí (1999) and Francis and Ramey (2005) have gained a great deal of attention lies in their implications for DSGE modeling. In a New Keynesian environment, price rigidities prevent that aggregate demand adjusts as fast as aggregate supply. Hence, when the degree of price rigidity is sufficiently large, hours worked may fall in the aftermath of a positive productivity shock (Galí and Rabanal, 2005). Furthermore, the fall in hours worked to neutral technology shocks implies that neutral technology shocks cannot be the main source of macroeconomic fluctuations — otherwise, hours worked would be countercyclical. Our results, however, do not provide strong support for this view. In contrast, the inflation rate drops significantly in response to neutral technology shocks. This suggests that price rigidities are rather moderate, causing only a small output gap.[19] Consequently, the median response of per-capita hours is greater than zero over the whole observation period.

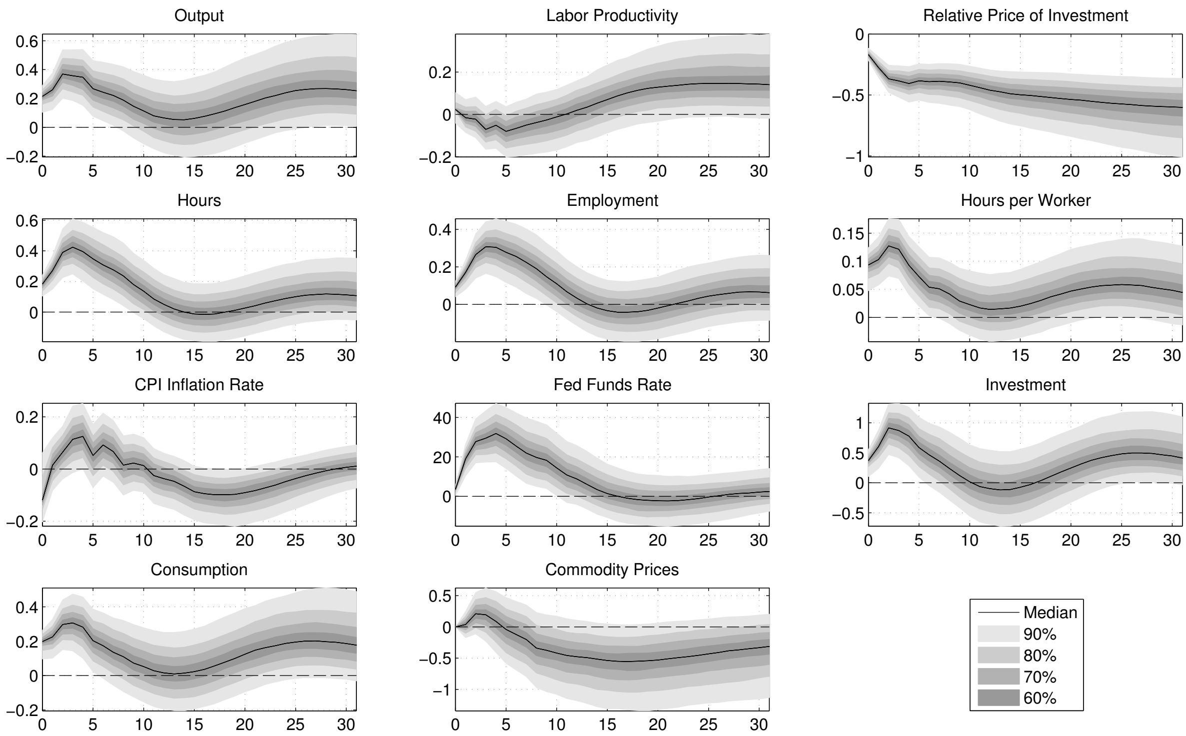
3.1.2 Investment-Specific Technology Shocks

Figure (5) displays the effects of an investment-specific technology shock. This shock leads to a sudden and permanent drop in the relative price of investment goods. We ob-

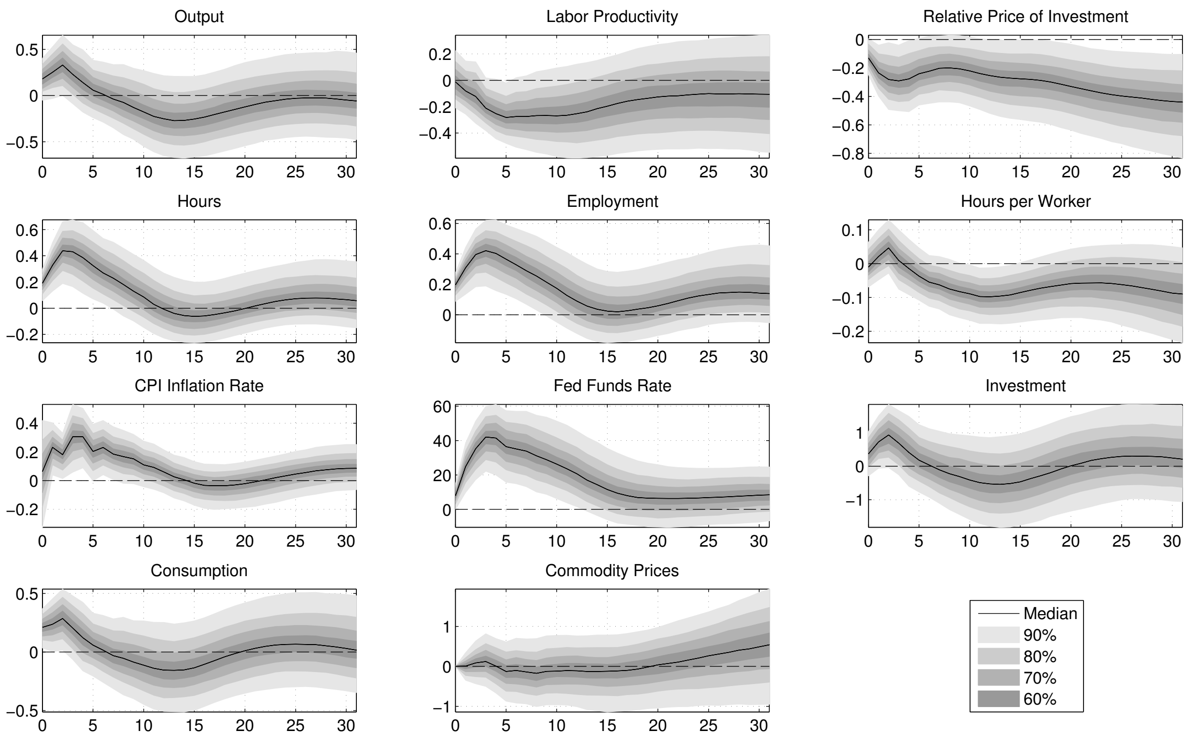
[17] Otherwise, the set-up is equivalent to the level specification.

[18] Using a trivariate VAR, Forni and Gambetti (2011) demonstrate that such a small model is not informationally sufficient. As in our SVAR, the impulse response of per-capita hours to technology shocks is hardly significantly different from zero when they include additional information.

[19] Dupor et al. (2009) draw a similar conclusion.



(a) Benchmark Specification



(b) Level Specification

Figure 5: The figure illustrates the impulse responses to an investment-specific technology shock.

serve that all variables (except labor productivity) move together in response to this type of disturbance. Their dynamic adjustment paths show a marked hump-shaped pattern, with peak effects occurring after 3-4 quarters. The impulse response of labor productivity remains insignificant for more than four years before eventually rising. This result illustrates that, on impact, the elasticity of per-capita hours is of the same magnitude as aggregate output (also here, most variation in labor input is due to adjustments along the extensive margin). Thus, investment-specific technology shocks seem far more important for the cyclical behavior of the labor market than neutral technology shocks. Against this background, it is a little surprising that investment-specific technology shocks have not received more attention as a driving force of labor market fluctuations. Two of the exceptions are De Bock (2007) and Faccini and Ortigueira (2010) who both study the implications of frictional labor markets in this context. Overall, we note that the estimated adjustment dynamics are virtually identical to the results of Altig et al. (2011).

The main reason why investment-specific technology shocks have not played a more important role in the literature has recently been pointed out by Justiniano et al. (2010). Their work explains why the standard RBC model fails to replicate the positive comovement among output, hours worked, and consumption over the business cycle. In such a frictionless environment, the marginal product of labor (MPL) equals the marginal rate of substitution (MRS) between consumption and leisure. Investment-specific technology shocks, however, have no direct impact on total factor productivity. This implies that consumption and leisure move in the same direction or, in other words, consumption and hours worked move in opposite directions. In the data, however, consumption and hours worked are positively correlated at business cycle frequencies. For this reason, investment-specific technology shocks were long considered not to be a main determinant of the business cycle (Barro and King, 1984). Neutral technology shocks, on the contrary, are able to match the positive comovement easily. Hence, neutral technology shocks attracted most attention in the early RBC literature (King and Rebelo, 1999).

Justiniano et al. (2010) also demonstrate how the standard RBC model can be reconciled with the empirical evidence. They suggest considering real frictions (or the like) that modify the relationship between the MPL and MRS. With variable capital utilization, for instance, investment-specific technology shocks may have a direct impact on the MPL. Consequently, consumption and per-capita hours do not necessarily move in opposite directions.[20] Moreover, in the presence of monopolistic competition, firms set prices as a mark-up over marginal costs. If the mark-up is time-variant (Rotemberg, 2008), it may drive a wedge between the MPL and the MRS. Besides, Justiniano et al. (2010) also consider habit formation in aggregate consumption (Abel, 1990). This modification alters the marginal utility of consumption and, thus, the functional form of the MRS. The previous section, however, has shown that external habit formation is inconsistent with the dynamics of the U.S. economy in response to neutral technology shocks.

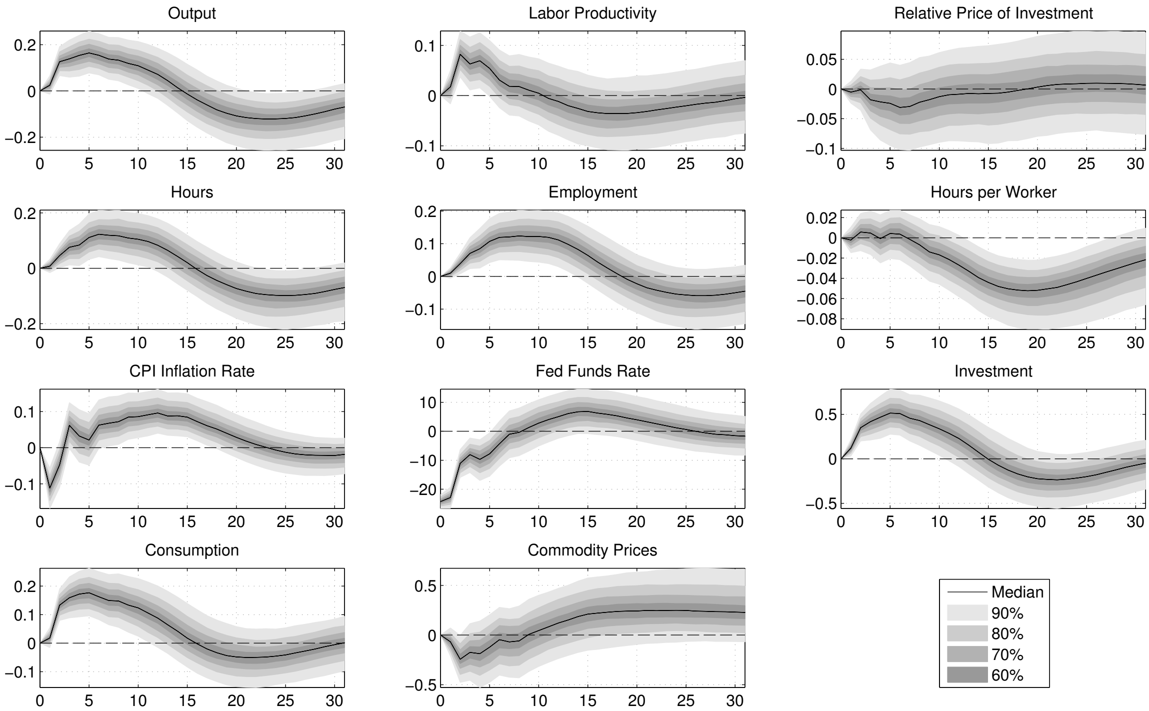
[20] Furlanetto and Seneca (2010) examine under which parameter values medium-scale DSGE models are able to generate a positive consumption response to investment-specific technology shocks.

3.1.3 Monetary Policy Shocks

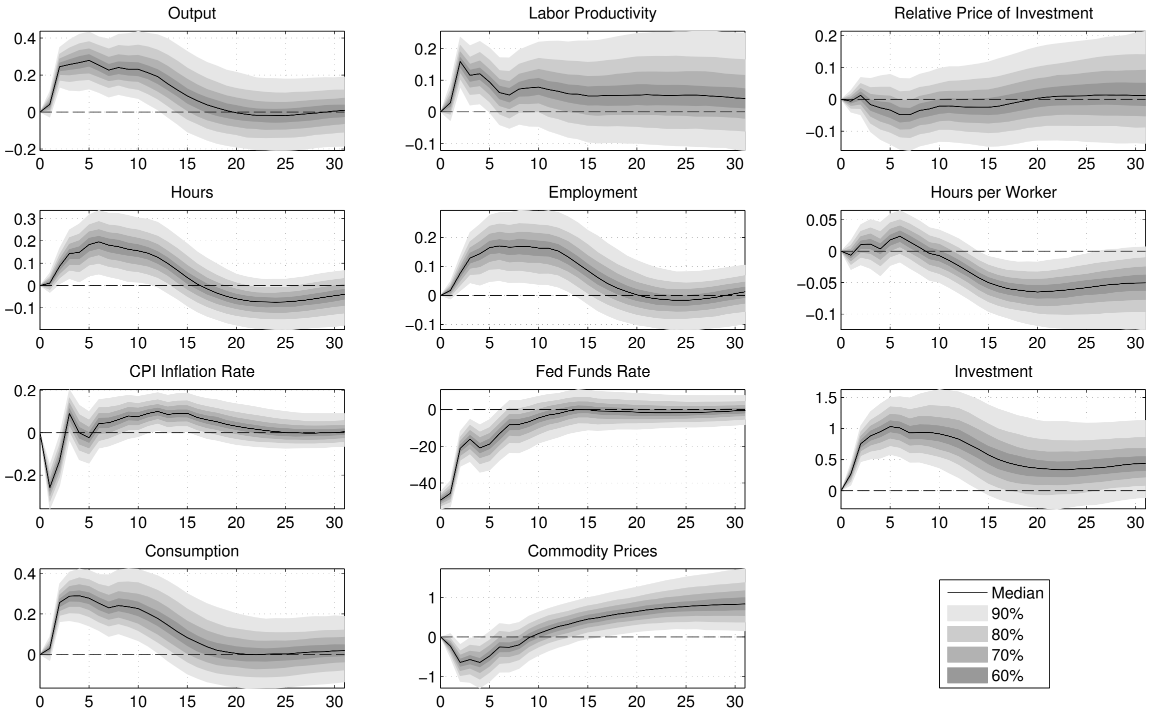
Figure (6) shows the responses to an expansionary shock in monetary policy. This shock represents a drop in the Federal Funds rate, due to an unexpected deviation from the Fed's linear policy rule. Our identifying assumptions imply that the shock has only a temporary effect. Nevertheless, the Federal Funds rate remains below its steady state level for more than seven quarters. In response to this, we observe that output, per-capita hours, employment, consumption, and investment rise gradually. Peak effects take place about 5-6 quarters after the monetary stimulus. At longer forecast horizons, the adjustment paths show a slight rebound. The response of the relative price of investment goods, on the other hand, is not significant. Overall, the shapes and elasticities of the responses are in line with the estimates by Ravn and Simonelli (2008). Only labor input indicators behave slightly different. Employment seems somewhat less elastic. Hours per worker even display a very mild downturn.

Consistent with Sims (1992), the impulse response of the inflation rate drops on impact, followed by a slow and persistent increase. The former observation is often referred to as the "price puzzle", the latter as "inflation persistence". The slow speed of aggregate consumer price adjustment in the aftermath of a monetary policy shock has led to the development of medium-scale DSGE models, including nominal rigidities and several other departures from the standard RBC economy (Christiano et al., 2005). However, estimated versions of these models suggest that the degree of price rigidity is much larger than can be supported by micro data (see, e.g., Bils and Klenow, 2004). Hence, various extensions have been proposed to rationalize inflation persistence; for instance, deep habits (Ravn et al., 2010), firm-specific capital (Altig et al., 2011), or rational inattention (Mackowiak and Wiederholt, 2011). Nevertheless, the temporary drop remains "puzzling". As explained by Eichenbaum (1992), it may be due to the presence of deflationary pressure already known to the Fed, but not captured by the information set of the VAR. However, given that the index of commodity prices is a forward-looking variable, our SVAR should be immune to this critique. According to our estimates, the inclusion of the commodity price index reduces the size of the drop slightly, but leaves the shape of the inflation response unchanged. This indicates that the "price puzzle" is a robust feature of the data.

Recently, Ravn et al. (2010) were able to match the shape of the inflation response to monetary policy shocks assuming "deep" habits over individual varieties of consumption goods (instead of aggregate consumption). In this set-up, mark-ups are endogenous. The authors show that, in response to a cut in the nominal interest rate, firms have an incentive to lower their mark-ups. The rationale behind this behavior is the following: When the expected value of future market shares is high, firms have large incentives to increase the stock of habits. Once the stock of habits has been built up, firms exploit the low price elasticity of demand. For this reason, we observe an initial drop in the inflation rate,



(a) Benchmark Specification



(b) Level Specification

Figure 6: The figure illustrates the impulse responses to a monetary policy shock.

followed by a protracted increase. Due to the presence of complementarities, the deep habits model requires — consistent with micro evidence — only a low to moderate degree of nominal rigidities to replicate the gradual response of the inflation rate. As pointed out by Justiniano et al. (2010), counter-cyclical mark-ups also help to replicate the positive comovement of output, hours worked, and consumption in response to investment-specific technology shocks. Hence, it seems promising to evaluate the deep habits model also in response to other structural shocks.[21]

Furthermore, we observe that an unexpected cut in the Federal Funds rate induces a slow, but persistent increase in commodity prices. The maximum impact does not occur until four to five years after the shock. In comparison to Anzuini et al. (2010), our estimated impulse response is much more gradual and resembles (qualitatively as well as quantitatively) the impulse response of the consumer prices index (i.e., the cumulative response of the inflation rate). In other words, the commodity price index shows no significant response in *real* terms.

3.1.4 Commodity Price Shocks

Figure (7) depicts the impulse responses to the identified commodity price shock.[22] We find that this shock triggers a temporary rise in the commodity price index, peaking shortly after the initial increase before slowly returning to its steady state level. Moreover, we observe a spike in the inflation rate, indicating that aggregate consumer prices are very flexible in response to commodity price shocks. In the following periods, the inflation rate declines sharply. The sudden surge in the inflation rate prompts the Fed to elevate the nominal interest rate for a protracted period (about 6-8 quarters).[23] Consequently, the inflation rate falls below normal about two years after the shock. We also note that the relative price of investment goods decreases slightly, but the effect disappears relatively quickly. The adjustment paths of output, per-capita hours, employment, hours per worker, consumption, and investment display significant U-shaped responses.

The estimated impulse responses of output and employment are consistent with the results of Blanchard and Galí (2010) — output and employment decline persistently after a lag of 3-5 quarters and reach a trough after about ten quarters.[24] Ordóñez et al. (2011) also find that energy price shocks cause upheaval in the labor market, albeit with a shorter lag. Rotemberg and Woodford (1996) emphasize that energy price shocks are often accompanied by upward movements in firms' mark-ups, thus exacerbating the economic downturn.

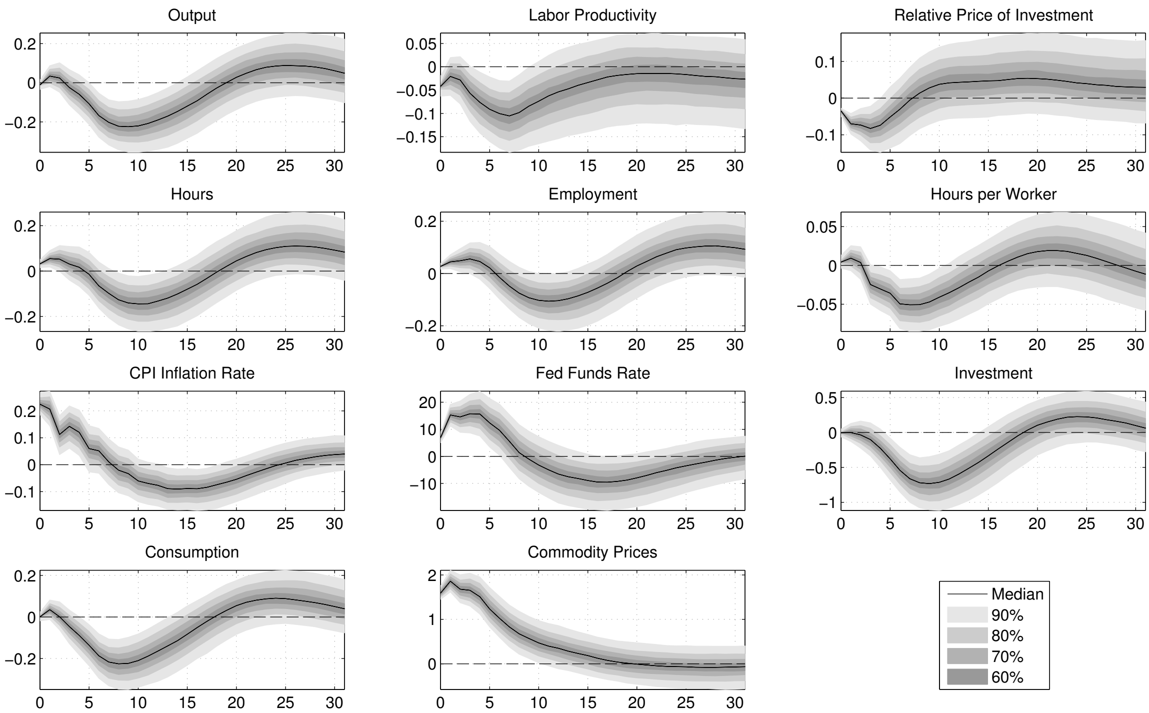
Furthermore, Bernanke et al. (1997) argue that a substantial part of the recessionary effects of oil price shocks is not due to the direct impact of higher producer prices, but rather due to the systematic contractionary response of the Federal Reserve. Their conclusion stems from a counterfactual exercise which presumes that the Federal Funds rate

[21]For instance, di Pace and Faccini (2011) examine the impact of neutral technology shocks in a RBC economy with deep habits and frictional labor markets.

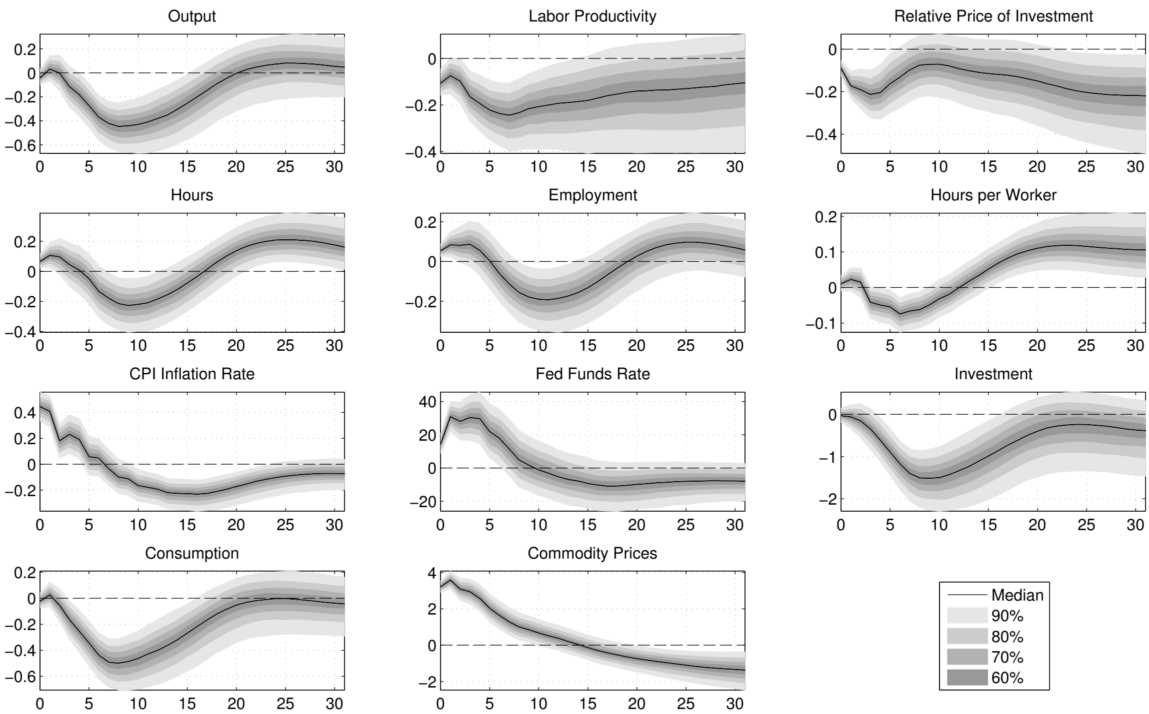
[22]Section 4.3 provides a robustness analysis on the impact of demand-driven commodity price changes.

[23]The peak response of the Federal Funds rate corresponds to 30 basis points. We infer this value from the *level* specification (see Figure 7, panel b).

[24]A detailed description of the transmission mechanism can be found in Kilian (2008).



(a) Benchmark Specification



(b) Level Specification

Figure 7: The figure illustrates the impulse responses to a commodity price shock.

is kept constant when the U.S. economy is hit by an unexpected increase in commodity prices. In the following, we perform the same counterfactual exercise using our benchmark model. Importantly, our benchmark model accounts for two issues emphasized by Hamilton and Herrera (2004). First, we use a large information set. This implies that the

Federal Reserve in our counterfactual exercise responds to eight (three) out of nine (four identified) shocks. Hence, we observe that the imputed movements in the Federal Funds rate deviate only moderately from the original series (see Figure 8). For this reason, we believe that our results are less prone to changing parameters due to the Lucas (1976) critique. Second, we use a lag length beyond one year.

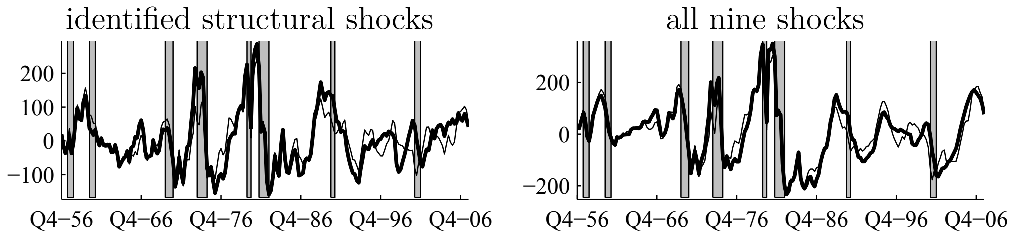


Figure 8: The left panel contrasts the time path of the Federal Funds rate predicted by the four identified structural shocks (bold line) with the same series in the absence of a monetary policy feedback rule (thin line). The right panel contrasts the Federal Funds rate predicted by all nine shocks (bold line) with the same series in the absence of a monetary policy feedback rule (thin line).

Figure (9) contrasts the impulse response functions of output, per-capita hours and the inflation rate in the benchmark specification (top panel) with the impulse responses under the counterfactual assumption (bottom panel). Indeed, we are unable to observe a significant downturn in output and per-capita hours if the Fed stayed passive. There is only an insignificant decline in output that occurs with a lag of about two years. The *initial* spike in the inflation rate, on the other hand, seems identical to the one estimated in the benchmark specification. *At medium horizons* (10-20 quarters), however, the counterfactual response cannot replicate the significant disinflationary rebound in CPI inflation. Thus, we conclude that the contractionary monetary policy feedback rule helped the Federal Reserve to achieve price stability in the *long run*, yet at the cost of a significant economic downturn in output and per-capita hours.[25]

In addition, Figure (10) illustrates the impulse responses of two CPI sub-indices; i.e. the so-called core inflation rate (all items less food and energy) and its counterpart (food and energy only). We observe that the spike in the headline inflation rate is mainly due to a sharp rise in food and energy prices. The core inflation rate, on the other hand, shows a lower — but still significant — and more persistent increase. This indicates that a little price rigidity at the level of intermediate goods may translate into persistent inflation movements in other sectors of the economy (Basu, 1995) — so-called second-round effects. We also find a marginally significant disinflationary rebound in both CPI sub-indices at medium horizons. Moreover, by repeating the above-described counterfactual exercise,

[25] Kilian and Lewis (2011), on the other hand, conclude that the combined direct and indirect effect on the U.S. economy has been negligible. Their counterfactual analysis, however, differs from the one presented here in that they remove energy prices from the Fed's information set.

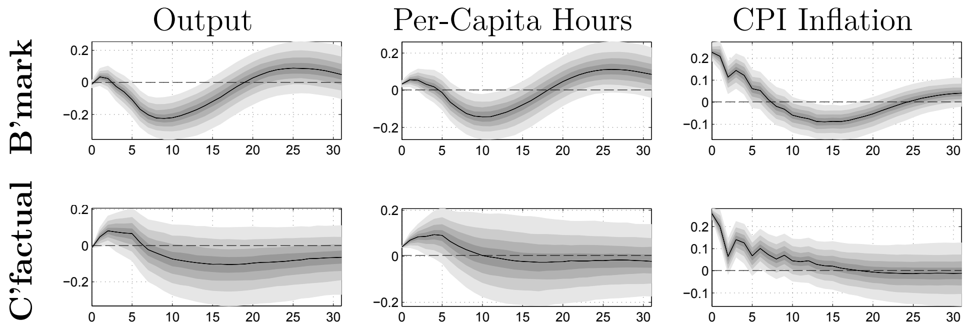


Figure 9: The top panel illustrates the responses of output, per-capita hours, and inflation to the estimated commodity price shock. The bottom panel illustrates the same responses when the Federal Funds rate — counterfactually — is assumed to be constant.

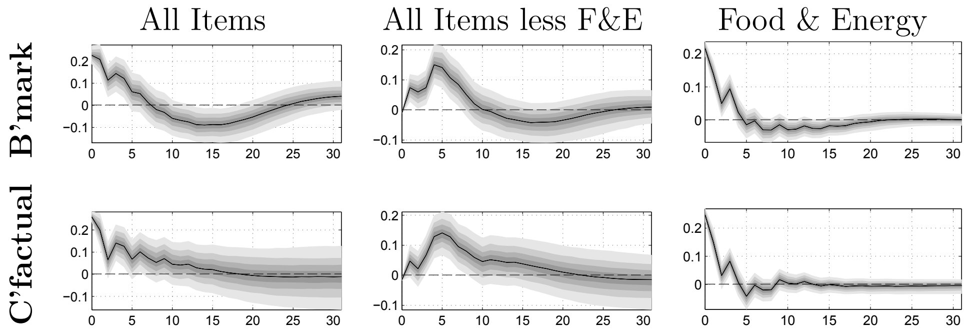


Figure 10: The figure illustrates the impulse responses of three CPI inflation measures to the identified commodity price shock; i.e., the "headline" inflation rate (all items), the "core" inflation rate (all items less food and energy), and the "food and energy" inflation rate. Due to limited data availability, the latter two responses are estimated using a slightly reduced sample period (1958Q2-2007Q4). The CPI "food and energy" is a weighted average of its components, using time varying weights (based on own calculations). All data are taken from FRED.

we notice that the initial increase in both sub-indices remained virtually unchanged if the Federal Reserve stayed passive. The disinflationary rebound, however, disappears in both impulse responses. Therefore, we conclude that the Fed's contractionary monetary policy feedback rule is unable to avoid second-round effects in the *short run*. Yet, it exhibits medium-run disinflationary effects which help the Federal Reserve to achieve price stability at *longer forecast horizons*.

3.2 Importance of the Structural Shocks

We now examine the relative importance of the four identified structural shocks for the variance of all variables included in our SVAR. First, we present the share of variation explained by each identified shock at different forecast horizons. However, as explained by Ravn and Simonelli (2008), these figures do not allow us to draw direct conclusions about the importance of these shocks at business cycle frequencies. Therefore, we also compute the variance decomposition at business cycle frequencies (8-32 quarters) following the method proposed by Altig et al. (2011). In addition, we present the historical decomposition of the four identified structural shocks for aggregate output in the postwar period.

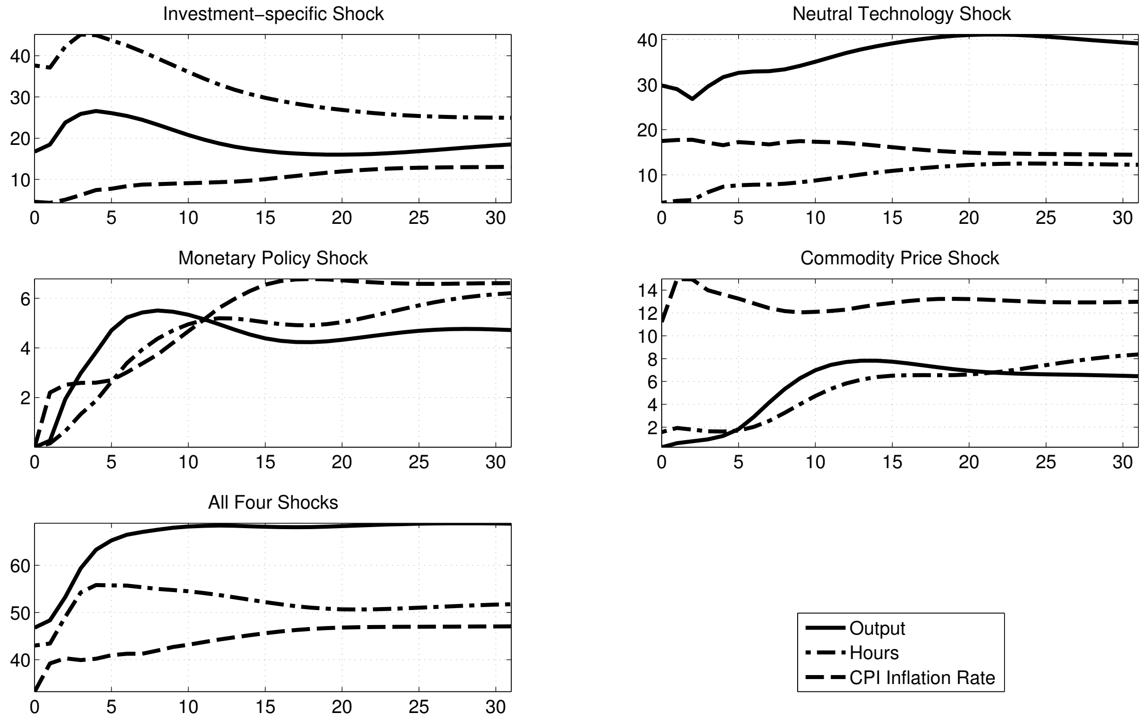


Figure 11: Forecast Error Variance Decomposition – The figure illustrates the forecast error variance decomposition in our benchmark specification.

3.2.1 Forecast Error Variance Decomposition

Figure (11) displays the forecast error variance decomposition of three key macroeconomic variables (output, per-capita hours, and the inflation rate) at different horizons. We observe that neutral technology shocks explain a large share of the variation in output, particularly at long forecast horizons. Investment-specific technology shocks are the main determinant of fluctuations in per-capita hours, and the second most important determinant of fluctuations in output. Commodity price shocks (together with neutral technology shocks) appear to be the primary driving force of movements in the inflation rate. Monetary policy shocks, on the other hand, explain only small shares of macroeconomic fluctuations. The joint explanatory power of all four shocks lies between 35% (inflation rate) and 48% (output) in the short run, and between 48% (inflation rate) and 68% (output) in the long run.

3.2.2 Variance Decomposition at Business Cycle Frequencies

We now investigate the variance decomposition at business cycle frequencies (see Table 3). The results suggest that commodity price shocks are a principal driving force of macro-

Table 3: Variance Decomposition at Business Cycle Frequencies

	investment specific	neutral tech	monetary policy	commodity prices	all four shocks
y_t	21 (12)	11 (7)	8 (5)	12 (8)	52 (14)
y_t/h_t	13 (9)	17 (9)	6 (4)	8 (6)	44 (14)
q_t	29 (15)	11 (7)	3 (3)	9 (6)	51 (16)
h_t	29 (14)	6 (6)	6 (5)	11 (8)	52 (14)
n_t	27 (14)	6 (6)	7 (5)	11 (8)	51 (14)
h_t/n_t	29 (13)	9 (6)	5 (4)	13 (8)	55 (14)
π_t	14 (10)	12 (8)	6 (4)	17 (9)	49 (14)
r_t	26 (14)	7 (6)	11 (6)	14 (8)	58 (14)
i_t	20 (12)	9 (6)	8 (5)	14 (8)	51 (13)
c_t	20 (11)	13 (8)	8 (5)	14 (8)	55 (14)
p_t	11 (8)	7 (6)	4 (3)	51 (15)	73 (14)

(a) Benchmark Specification

	investment specific	neutral tech	monetary policy	commodity prices	all four shocks
y_t	17 (10)	14 (9)	7 (4)	17 (8)	55 (11)
y_t/h_t	19 (10)	16 (10)	5 (3)	11 (7)	51 (12)
q_t	18 (14)	10 (6)	3 (2)	13 (7)	43 (14)
h_t	20 (12)	9 (8)	6 (4)	14 (8)	50 (13)
n_t	23 (13)	9 (8)	6 (4)	14 (8)	52 (13)
h_t/n_t	14 (9)	9 (7)	5 (4)	17 (8)	45 (11)
π_t	17 (11)	9 (7)	5 (4)	25 (10)	57 (11)
r_t	20 (12)	8 (7)	13 (5)	17 (9)	58 (13)
i_t	16 (10)	11 (8)	8 (5)	17 (8)	53 (11)
c_t	15 (9)	14 (8)	8 (5)	20 (9)	57 (11)
p_t	7 (4)	5 (4)	6 (4)	60 (11)	77 (10)

(b) Level Specification

Notes: The table displays the decomposition of variance at business cycle frequencies based on estimated spectral densities (following Altig et al., 2011). Numbers are means of point estimates across bootstrap simulations, numbers in parentheses are the corresponding standard deviations.

economic fluctuations. In particular, we find that commodity price shocks explain a large share of cyclical movements in inflation. The commodity price shock also turns out to be a very important determinant of cyclical fluctuations in many other macroeconomic variables (e.g., Federal Funds rate, investment, or consumption), second only to investment-specific technology shocks. The neutral technology shock explains only a considerable share of the variation in labor productivity — the endogenous variable in the equation that identifies the neutral technology shock. The monetary policy shock seems even less relevant. Two observations provide support for our identification strategy. On the one hand, only 11% of the changes in the nominal interest rate are due to the unexpected shock. This implies that the Fed's monetary policy has followed a rule-based approach over our sample period. On the other hand, more than 50% of the cyclical variability in the commodity price index is explained by the commodity price shock itself. This result indicates that the contemporaneous exogeneity assumption is a reasonable identifying restriction.

The importance of investment-specific technology shocks is in line with the results of several recent SVAR studies by Fisher (2006), Ravn and Simonelli (2008), Canova et al. (2010), and Altig et al. (2011).[26] Altogether, the four identified shocks account for 49%-73% of business cycle volatility in the data. At first glance, however, it seems surprising that neutral technology shocks do not explain larger shares at business cycle frequencies. Therefore, we analyze also the explanatory power of neutral technology shocks across the whole spectrum (Figure 12). Indeed, we find that neutral technology shocks play a very important role in explaining macroeconomic fluctuations (particularly, output, labor productivity, and consumption), but at low frequencies.

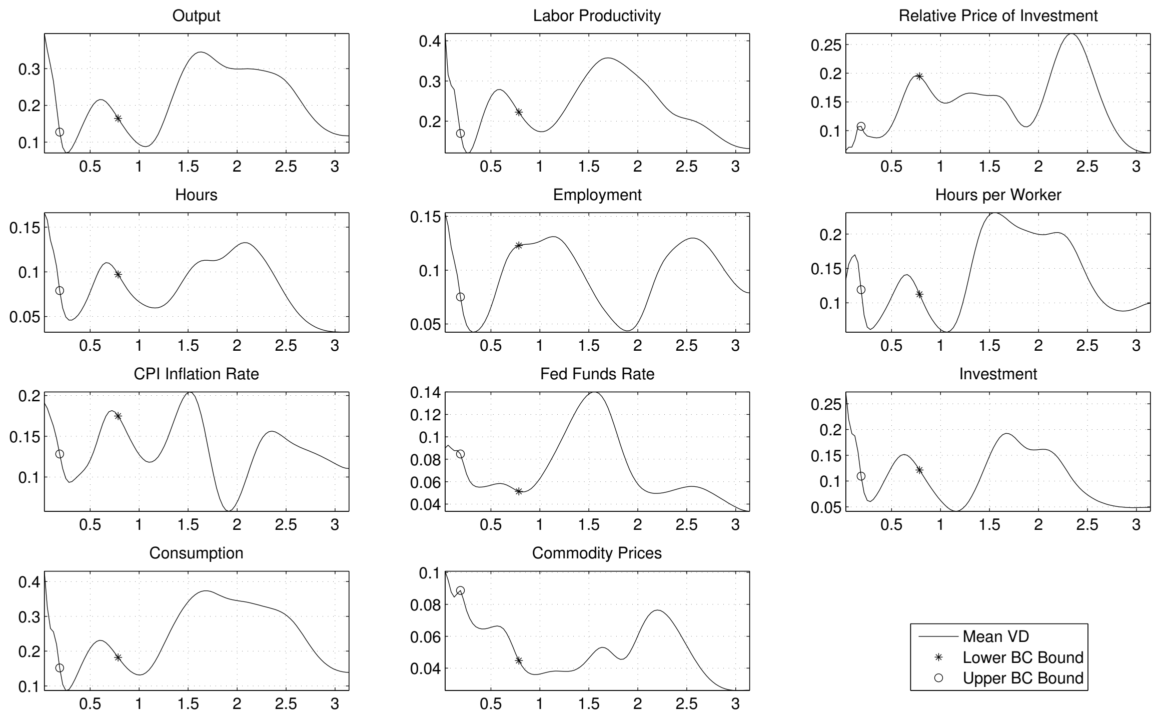


Figure 12: Variance Decomposition at the Frequency Domain – The figure illustrates the explanatory power of neutral technology shocks across the whole spectrum in our benchmark specification.

3.2.3 Historical Decomposition of Shocks

Figure (13) presents the historical decomposition of shocks for aggregate output. When all four identified structural shocks are considered, we observe that our SVAR model is

[26]Smets and Wouters (2007) as well as Mumtaz and Zanetti (2011) draw the opposite conclusion from a Bayesian VAR model using a data set that includes consumer durables in consumption (and not in investment). Schmitt-Grohé and Uribe (2011) argue that a common stochastic trend in neutral and investment-specific technology is the main driving force of the business cycle.

able to replicate the cyclical behavior of output remarkably well. There are only two episodes in U.S. postwar history that exhibit a noticeable tracking error. The model explains neither the short recession in the late 1960s, nor the depth of the recession after the burst of the so-called dot-com bubble.

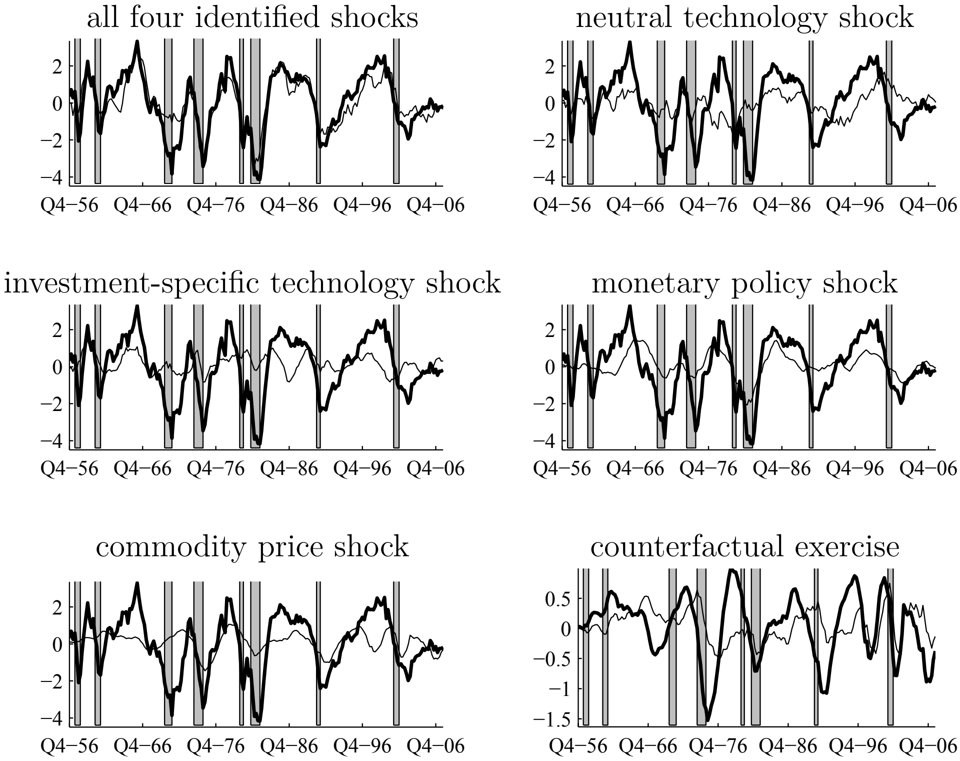


Figure 13: The figure illustrates the historical decomposition of the four identified structural shocks for output. The bold line represents the bandpass filtered data, the thin line represents the time series predicted by the respective shock(s). In addition, the counterfactual exercise contrast the output series predicted by the commodity price shock (bold line) with the output series predicted by the commodity price shock in the absence of the monetary policy feedback rule (thin line).

We also investigate the time series elicited by the four individual shocks. The graphs illustrate that their contribution varies considerably across different episodes in the U.S. postwar period. In line with our previous results, we are unable to find a systematic relationship between neutral productivity shocks and fluctuations in aggregate output at business cycle frequencies. Neutral technology shocks seem rather important at low frequencies. For example, neutral productivity shocks suggest a deep recession between 1976 and 1983, reflecting the productivity slowdown in that period, and two long-lasting economic booms — the first in the mid 1980s and the second in the late 1990s. Consistent with Greenwood et al. (1997), we find that investment-specific technology growth was particularly strong during the productivity slowdown of other factors in the 1970s. Furthermore, investment-specific technology shocks appear to be a principal driving force of the 1960-61 recession and the following economic expansion, the 1973-75 recession, the early 1980s recession as well as of the subsequent recovery. Monetary policy shocks, on the other hand, have played a role in the late 1960s recession, the double-dip in the early 1980s as well as in the subsequent recovery. Commodity price shocks contribute most to the high degree of macroeconomic volatility in the 1970s, particularly during and after the first OPEC oil crisis. In addition, commodity price shocks are also an important determinant of the double-dip in the early 1980s, the economic boom in the early 1990s, and the short early 2000s recession. The last graph evaluates the impact of the Fed's response to commodity price shocks. Therefore, we examine the cyclical movements of aggregate output in the absence of the contractionary monetary policy feedback rule.[27] Interestingly, we observe that for a long sub-period of our sample the monetary policy feedback rule — particularly the contractionary response during the first OPEC oil crisis and the subsequent monetary easing — seems to have amplified the output fluctuations caused by unexpected changes in commodity prices. Consistent with the estimated impulse response function, the counterfactual time series seems to lag the estimated path by about one year (see the delayed decline during the first OPEC oil crises and after the early 1990s recession). In other words, our SVAR indicates that a contractionary monetary policy feedback rule may help the Federal Reserve to achieve price stability at longer forecast horizons, yet at the cost of output destabilization.

4 Robustness Analysis

The following section presents a number of robustness checks. We investigate the sensitivity of our results to the data treatment, the choice of the lag length, the selected sample period, the identification of the commodity price shock, the usage of the Thomson Reuters (2010) Continuous Commodity Index,[28] and the inclusion of an external demand shock (Abbritti and Weber, 2010). We demonstrate that the results of our benchmark specification are robust across alternative model versions.

4.1 Data Treatment

4.1.1 Bandpass Filter vs. Level Specification

The (b) panels of Figures (3), (5)-(7) and Table (4), respectively, display the impulse responses and the business cycle variance decomposition when we estimate the SVAR in levels.[29] We observe that all major conclusions drawn from the benchmark specification survive this type of test. Even the response of per-capita hours to neutral technology shocks remains virtually unchanged (see also Section 3.1.1). The only notable difference

[27] See Section 3.1.4 for details and motivation of this counterfactual exercise.
[28] See also Footnote (4).
[29] See Footnote (16) for a definition of the level specification.

Table 4: Robustness

	investment specific	neutral tech	monetary policy	commodity prices	all four shocks
	Benchmark Specification				
y_t	21 (12)	11 (7)	8 (5)	12 (8)	52 (14)
h_t	29 (14)	6 (6)	6 (5)	11 (8)	52 (14)
π_t	14 (10)	12 (8)	6 (4)	17 (9)	49 (14)
	Level Specification				
y_t	17 (10)	14 (9)	7 (4)	17 (8)	55 (11)
h_t	20 (12)	9 (8)	6 (4)	14 (8)	50 (13)
π_t	17 (11)	9 (7)	5 (4)	25 (10)	57 (11)
	Difference Specification				
y_t	15 (10)	15 (9)	9 (5)	15 (7)	53 (12)
h_t	21 (12)	11 (9)	7 (5)	12 (7)	51 (13)
π_t	15 (10)	9 (6)	6 (4)	24 (9)	54 (12)
	Dummy Specification				
y_t	17 (10)	13 (8)	8 (5)	16 (8)	54 (11)
h_t	21 (13)	7 (6)	6 (4)	15 (8)	50 (13)
π_t	12 (9)	8 (6)	6 (4)	27 (10)	53 (11)
	Francis and Ramey (2009) Hours				
y_t	18 (11)	15 (9)	6 (4)	11 (6)	50 (12)
h_t	22 (13)	9 (8)	6 (4)	9 (6)	46 (13)
π_t	19 (13)	8 (7)	4 (3)	19 (9)	50 (12)
	Commodity Prices in First Differences				
y_t	18 (11)	14 (8)	8 (5)	10 (6)	50 (14)
h_t	28 (13)	6 (5)	6 (5)	9 (6)	50 (14)
π_t	12 (9)	8 (6)	7 (5)	17 (9)	44 (15)
	Benchmark Specification with 4 Lags				
y_t	15 (9)	13 (8)	6 (4)	11 (7)	45 (13)
h_t	21 (12)	8 (6)	5 (4)	10 (7)	44 (13)
π_t	7 (5)	12 (7)	7 (4)	20 (10)	45 (13)
	Pre-Volcker Period (1959Q1-1979Q2)				
y_t	15 (11)	13 (10)	12 (7)	16 (10)	57 (14)
h_t	14 (11)	11 (8)	13 (7)	17 (11)	55 (14)
π_t	15 (8)	10 (8)	12 (7)	17 (10)	55 (13)
	Post-Volcker Period (1980Q1-2007Q4)				
y_t	23 (14)	8 (7)	8 (5)	9 (7)	48 (14)
h_t	23 (16)	8 (7)	7 (5)	9 (8)	47 (15)
π_t	18 (12)	7 (6)	6 (4)	13 (8)	44 (13)
	External Demand Shock				
y_t	20 (12)	11 (7)	7 (5)	10 (7)	$54^{\dagger}$ (14)
h_t	28 (15)	7 (6)	6 (5)	10 (7)	$57^{\dagger}$ (14)
π_t	16 (10)	11 (9)	5 (4)	13 (8)	$51^{\dagger}$ (14)
	Non-Predetermined Commodity Prices				
y_t	18 (11)	12 (8)	7 (5)	12 (9)	49 (15)
h_t	27 (14)	7 (6)	6 (5)	12 (9)	52 (15)
π_t	15 (10)	11 (8)	5 (4)	17 (10)	48 (15)
	Thomson Reuters Continuous Commodity Index				
y_t	14 (9)	9 (6)	5 (3)	22 (10)	50 (13)
h_t	14 (10)	7 (6)	4 (3)	20 (10)	45 (13)
π_t	11 (7)	8 (6)	3 (3)	31 (12)	53 (12)

Notes: The table displays the decomposition of variance at business cycle frequencies based on estimated spectral densities (following Altig et al., 2011). Numbers are means of point estimates across bootstrap simulations, numbers in parentheses are the corresponding standard deviations. The "external demand shock" specification includes five shocks in total (denoted by a dag symbol †).

between these two specifications is that the cyclical variance decomposition statistics of the commodity price shock are higher in the level specification, but the cyclical variance decomposition statistics of the investment-specific technology shock are higher in the benchmark specification. Altogether, these results indicate that the low-frequency bias becomes less important when the information set is sufficiently large. Furthermore, the remarkable resemblance of the impulse responses suggests that bandpass filtering the data prior to estimation does not remove information necessary to identify the shocks using long-run restrictions (Gospodinov et al., 2011).

4.1.2 Treatment of the Hours Series

In addition, Table (4) provides the cyclical variance decomposition statistics of output, per-capita hours, and the inflation rate under different model specifications. The figures confirm that our findings are robust to different filtering methods (differences, dummies, including the corresponding Francis and Ramey (2009) hours time series in the level specification).

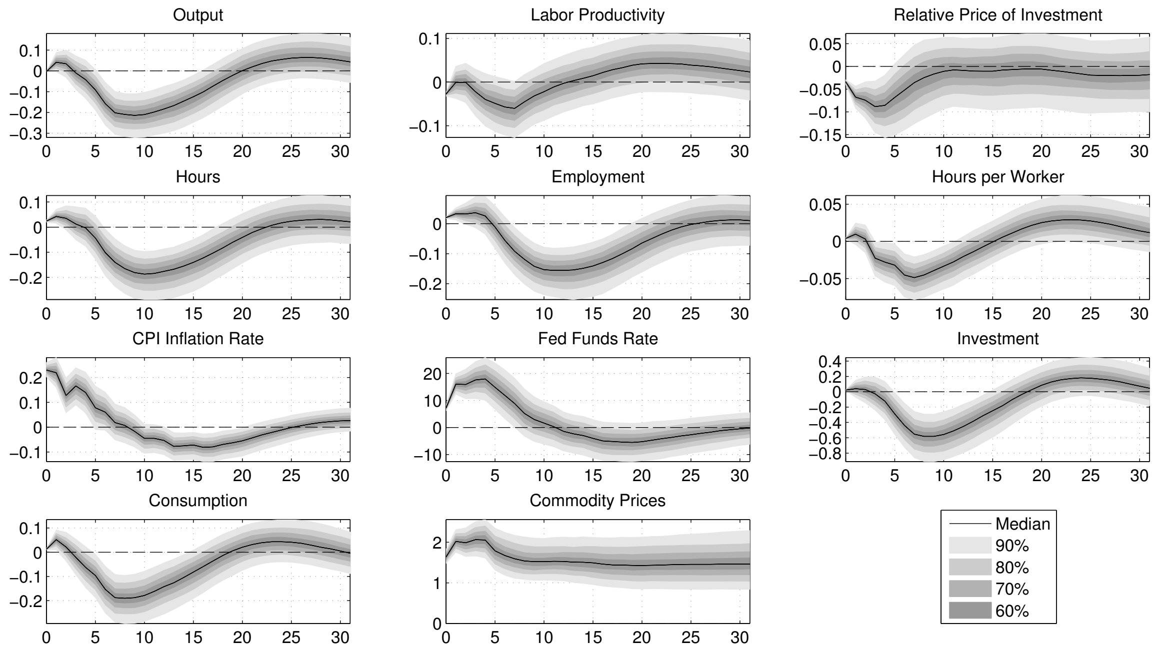


Figure 14: The figure illustrates the impulse responses to a commodity price shock when we first difference the natural logarithm of the commodity price index and then apply the one-sided bandpass filter.

4.1.3 Treatment of the Commodity Price Index

Figure (14) shows the impulse responses when the commodity price index is differenced prior to applying the one-sided bandpass filter. This implies that we now identify a *permanent* shock to the level of the commodity price index. We observe that the shapes of the impulse responses are almost identical to the benchmark specification. The most interesting difference is that the response of labor productivity is no longer significant, indicating that the elasticities of output and per-capita hours are of the same magnitude.

Also the business cycle variance decomposition statistics (see Table 4) are very similar to the benchmark specification.

4.2 Lag Length and Subsample Stability

The present section investigates whether the chosen lag length has any impact on our results. For this purpose, we reduce the number of lags to $M = 4$. Table (4) shows that, in this case, the investment-specific technology shock becomes less important, but remains the principal driving force of output and per-capita hours over the business cycle. This indicates that our SVAR may suffer from "truncation bias" (Erceg et al., 2005) when the VAR order is insufficiently short. Besides, we are unable to note any significant difference in the results.

Furthermore, we examine the subsample stability of our benchmark specification. Figure (15) illustrates the impulse responses to the identified commodity price shock before and after the appointment of Paul Volcker as chairman of the Board of Governors in August 1979. For this exercise — due to the smaller number of observations — we reduce the VAR order to $M = 3$. Note that, when plotting these graphs, we normalize the standard deviation of the commodity price shock in both sub-periods to the one measured in the full sample. We observe that output and the (core) inflation rate respond less elastic in the late sub-sample, but remain statistically significant at the 10% level. This result is in line with Blanchard and Galí (2010),[30] who attribute the milder response in the late subsample to (a) the smaller share of oil in production, (b) the decline in real wage rigidity, and (c) improvements in monetary policy.[31]

In contrast to their study, our SVAR explicitly controls for the decreasing share of oil in production (by using a broad commodity price index with time-varying weights) and identifies neutral and investment-specific technology shocks. Our result of smaller second-round effects is consistent with their view of a decrease in real wage rigidity. Moreover, we find evidence in favor of increased credibility of monetary policy. In the pre-Volcker period, we notice that the Federal Funds rate stays above its steady state level for about five quarters. Following the initial rise, the Federal Reserve reduces the nominal interest rate and keeps it below its long-run mean for the next ten quarters. This pattern is known as stop-and-go monetary policy.[32] In line with the conventional wisdom, we find no evidence for stop-and-go monetary policy in the post-Volcker period. The contractionary response in the post-Volcker period seems to be driven by the statistically

[30] Note that our sub-sample periods are not exactly identical to the ones chosen by Blanchard and Galí (2010). Given our 9-dimensional SVAR, the number of degrees of freedom is not sufficient in order to estimate the model accordingly.

[31] We also confirm their conclusion that the size of the shock in the post-Volcker period is larger than in the pre-Volcker period. This implies that the "Great Moderation" is *not due* to smaller commodity price shocks.

[32] The results of Evans and Fisher (2011) suggest that the stop-and-go pattern in the Federal Funds rate is triggered by oil price shocks, while the significant hike is due to changes in prices of other commodities.

significant hump-shape in the core inflation rate. Furthermore, consistent with the muted impulse responses, Table (4) shows that the explanatory power of the commodity price shock is somewhat lower in the post-Volcker period.[33]

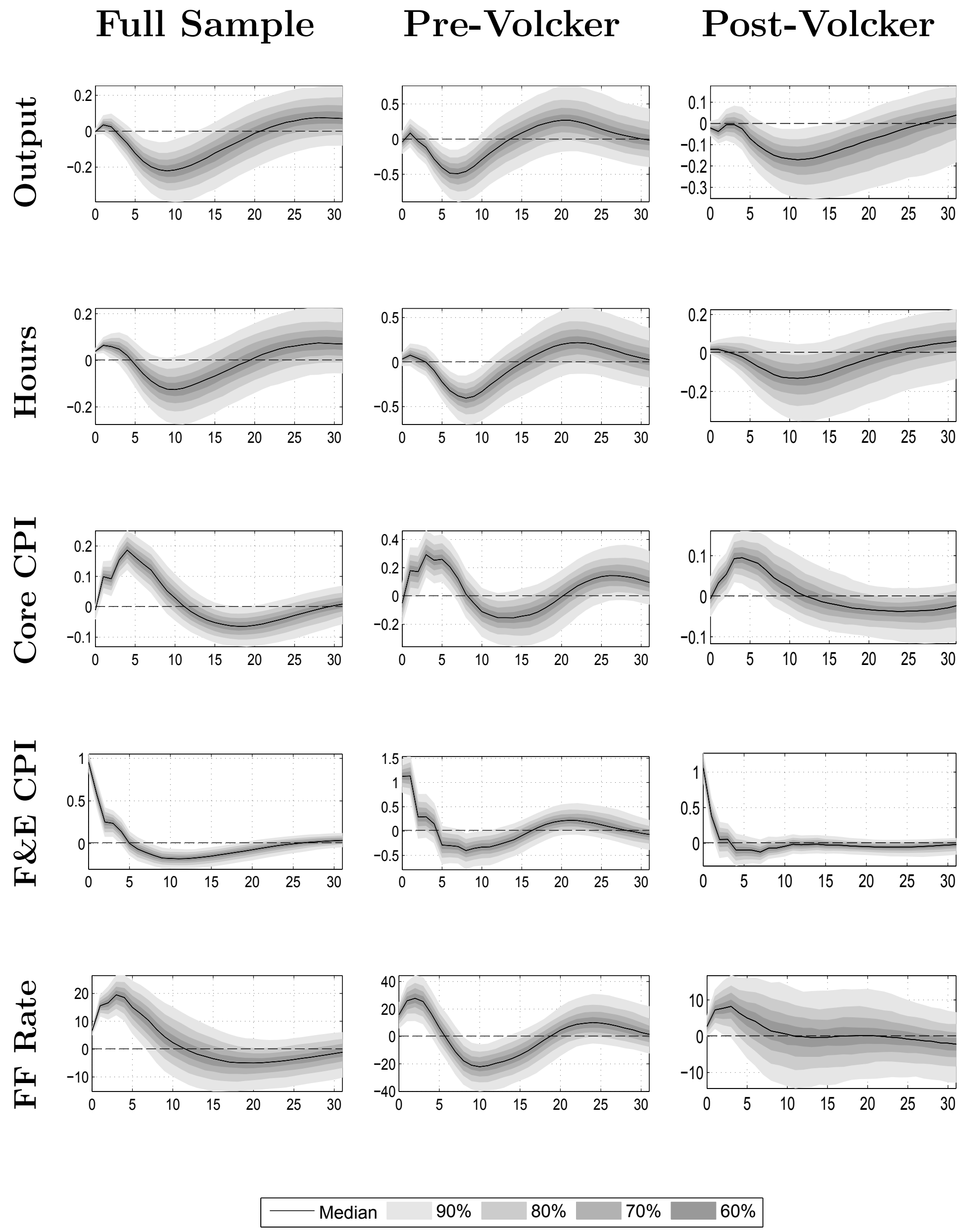


Figure 15: Subsample Stability – The figure contrasts the impulse responses to a commodity price shock in the "full" sample (1959Q1-2007Q4), the pre-Volcker period (1959Q1-1979Q2), and in the post-Volcker period (1980Q1-2007Q4). All three models are estimated with three lags ($M = 3$). In order to facilitate comparability across sub-samples, the standard deviation of the commodity price shock is normalized to the standard deviation over the full sample.

[33] Also note that the investment-specific technology shock is somewhat less important when we exclude the late 1990s Internet boom from our sample.

4.3 External Demand

The present identification procedure of the commodity price shock is unable to distinguish between supply- and demand-driven innovations. However, the assumption that commodity price shocks are contemporaneously exogenous to U.S. macroeconomic aggregates seems more defensible in the case of supply shocks (e.g., political strife in the Middle East) than in the case of demand shocks. Therefore, we extend our SVAR by adding a variable that captures variations in global demand for commodity goods. In particular, we choose to include the natural log of the ratio of real exports to real imports of goods and services (see Table 6 in Appendix 2.A). Based on this series, we identify an external demand shock using short-run restrictions. Following Abbritti and Weber (2010), we assume that the process for the real export/import ratio is independent of the current realizations of all other variables but the commodity price index.[34]

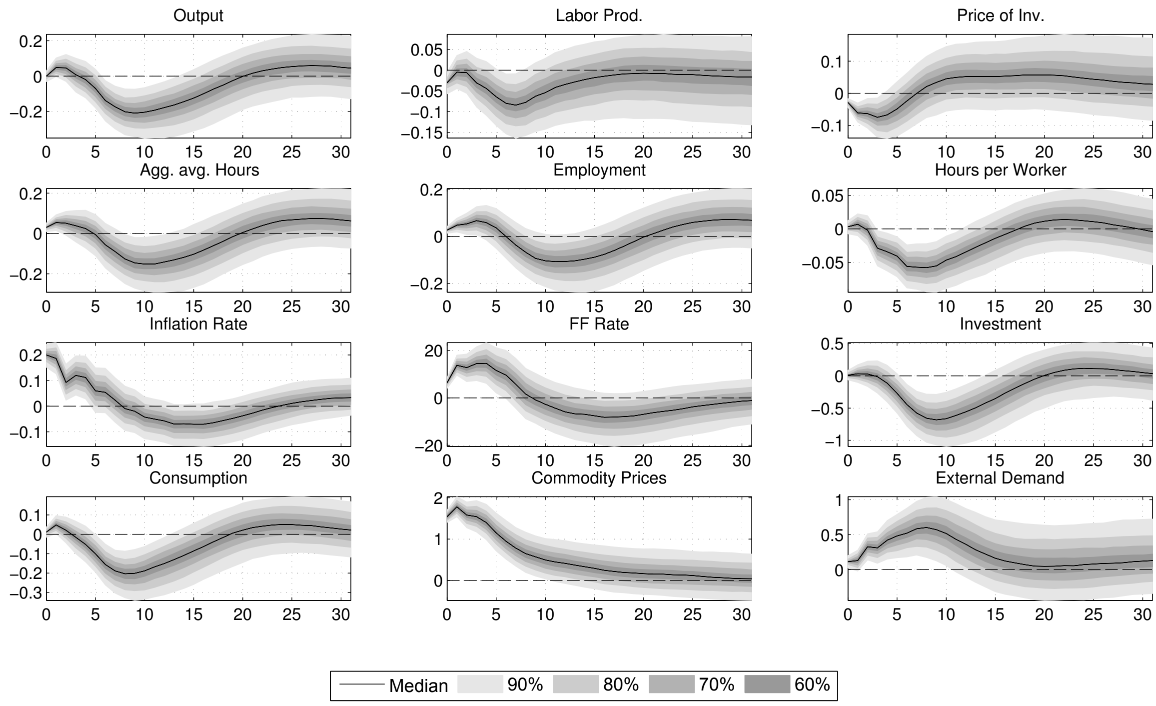
The impulse responses generated by the four remaining shocks, particularly by the commodity price shock (Figure 16a), remain virtually unchanged when we control for unexpected movements in external demand. In addition, the variance decomposition statistics at business cycle frequencies are remarkably robust (Table 4). We note only a mild reduction (4 percentage points) in the explanatory power of the commodity price shock with respect to cyclical movements in the inflation rate.[35] Figure (16b) illustrates the effects of the identified external demand shock. This shock represents a temporary but persistent rise in the real exports/imports ratio. We observe that the external demand shock causes a hump-shaped increase in the commodity price index, representing commodity price changes due to heightened global demand. Except for investment and consumption, all other variables show barely significant responses, which may be attributed to the fact that the U.S. is a relatively closed economy. Besides, the external demand shock is unable to explain significant shares in the business cycle variance of any variable but the real export/import ratio.

4.4 Non-Predetermined Commodity Prices

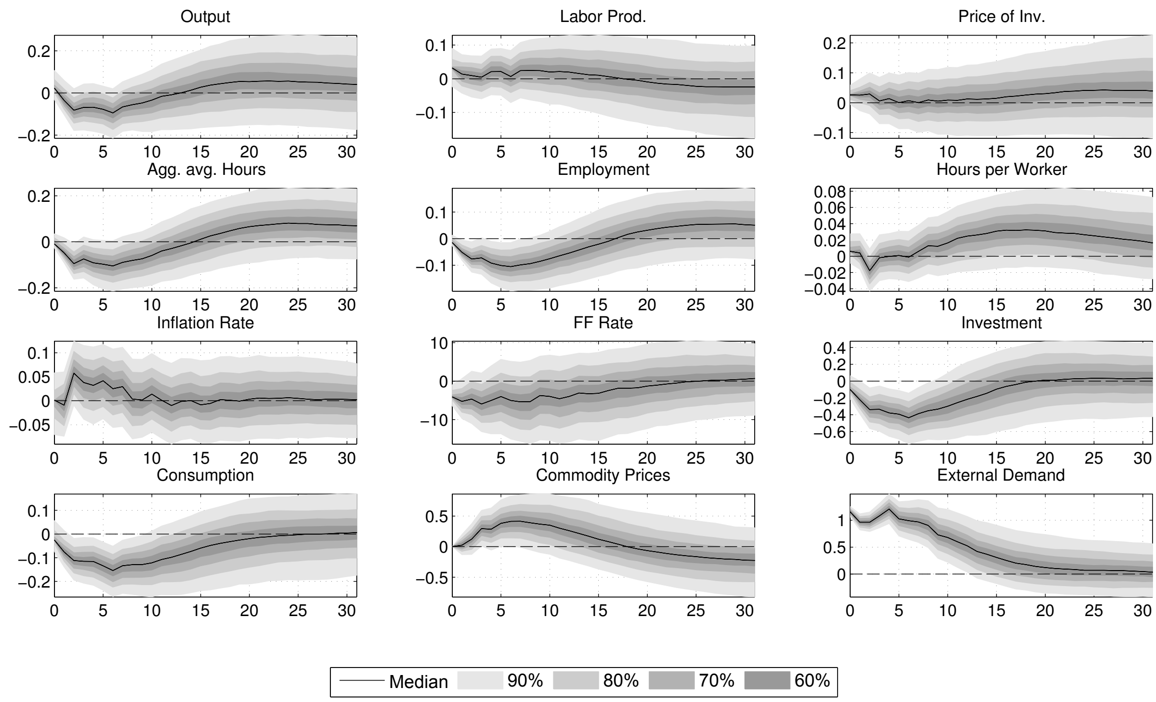
Our benchmark specification assumes that no single shock (but the commodity price shock itself) has an impact on the contemporaneous value of the commodity price index. Rotemberg and Woodford (1996) proposed this assumption to identify nominal oil price shocks. Given that the broad commodity price index may behave differently from the nominal oil price (Alquist et al., 2011), the current section examines the robustness of our

[34]Note that the idea to disentangle commodity demand and supply shocks is due to Kilian (2009). Baumeister and Peersman (2012b) identify oil supply vs. oil demand shocks using a sign restriction approach.

[35]Alternatively, we have used the "rest of the world" GDP index (1972Q1-2006Q4) by Enders et al. (2011), the "global economic activity" index (1968Q1-2007Q4) by Kilian (2009) and the "world industrial production" index (1948Q1-2007Q4) by Baumeister and Peersman (2012a). Our (subsample) tests indicate that all indices yield similar results.



(a) Commodity Price Shock



(b) External Demand Shock

Figure 16: The figure illustrates the impulse responses when we include external demand shocks.

identification strategy.

Therefore, we relax the contemporaneous exogeneity assumption to allow for immediate responses in the commodity price index to unexpected changes in two main indicators of the U.S. economy (labor productivity growth and per-capita hours). This procedure is similar to the procedure used by Blanchard and Galí (2010), who have explored the consequences of an alternative recursive ordering of the variables. However, as our SVAR is overidentified, we are able to eliminate these two identifying assumptions without imposing a new one.

Consequently, the identified commodity price shock ϵ_t^p is now obtained by estimating:

$$p_t = \alpha^p - \beta_{a,0}^p \Delta a_t - \beta_{h,0}^p h_t + \sum_{j=1}^{M} \beta_{x,j}^p x_{t-j} + \epsilon_t^p \tag{8}$$

Since ϵ_t^p may be correlated with Δa_t (via Equation 5) and h_t (via Equation 7), we estimate Equation (8) with 2SLS. The set of instruments includes a constant and the vector $[\Delta q_{t-j}, \Delta a_{t-j}, z_{t-j}, r_{t-j}, p_{t-j}]_{j=1}^{M+1}$.[36]

In line with Blanchard and Galí (2010), we find that the estimated results are remarkably similar under the two alternative identification schemes (see Figure 17). In particular, the CPI inflation rate behaves almost identical to our benchmark specification. Also the median response of the Federal Funds rate matches the benchmark estimate closely – even though the confidence bands are somewhat wider. Moreover, Table (4) shows that the business cycle variance decomposition statistics are very robust.

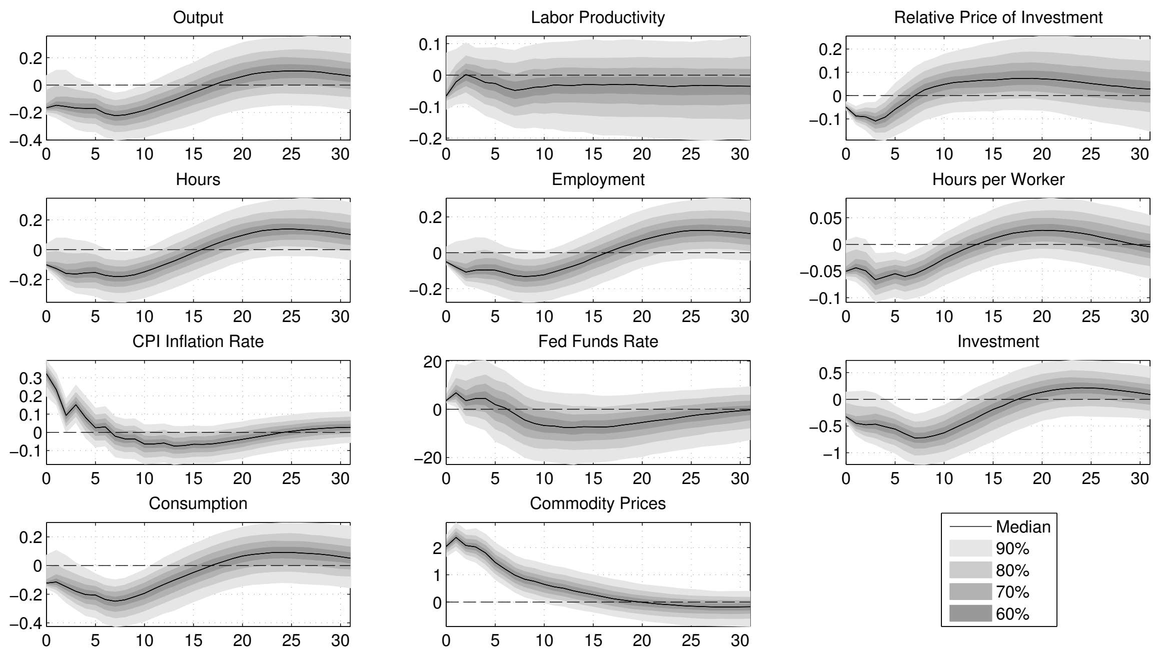


Figure 17: The figure illustrates the impulse responses to a commodity price shock when commodity prices may depend on the current values of labor productivity and per-capita hours.

[36]Note that the parameters are overidentified, given that the number of instruments exceeds the number of parameters. Using an overidentifying restrictions test (Sargan, 1964) we are unable to reject this specification at the 10% significance level.

4.5 Thomson Reuters Continuous Commodity Index

Figure (18) shows the impulse responses when the Thomson Reuters (2010) Continuous Commodity Index is used for the estimation of our SVAR model. In contrast to the benchmark specification, the alternative commodity price index is more persistent with peak effects occurring about one year after the initial increase. Nevertheless, the *relative* impulse responses appear very similar to the benchmark specification. Only the increase in the CPI inflation rate seems to be more long-lived. After the initial increase, we observe that the inflation rate remains elevated for more than one year before eventually falling back to normal. We also note that the confidence bands are tighter. Consequently, the commodity price shock now becomes significantly more important in terms of the cyclical variance decomposition statistics (see Table 4), particularly with respect to inflation. The investment-specific technology shock, on the other hand, loses some of its explanatory power.

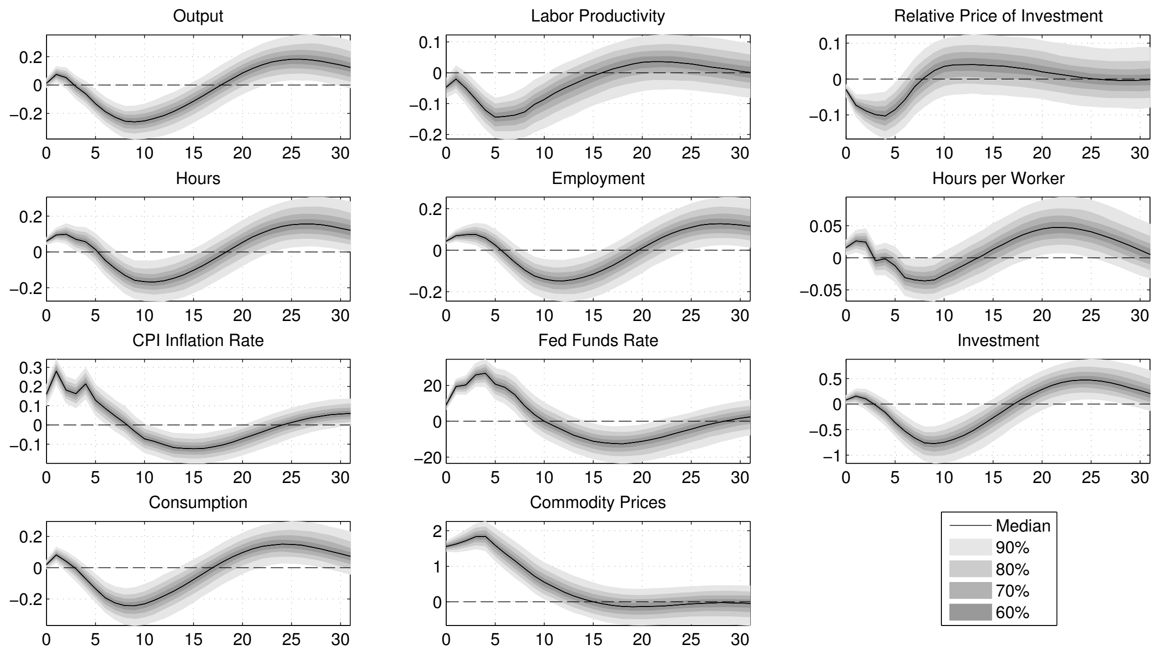


Figure 18: The figure illustrates the impulse responses to a commodity price shock when commodity prices are represented by the Thomson Reuters Continuous Commodity Index.

5 Summary and Conclusions

This chapter evaluates the importance of commodity price shocks in the U.S. business cycle. Therefore, we extend the standard set of identified shocks to include unexpected changes in commodity prices. The resulting SVAR shows that commodity price shocks are a very important driving force of macroeconomic fluctuations, second only to investment-specific technology shocks. In particular, commodity price shocks explain a large share of cyclical movements in inflation.

The impulse response analysis shows that commodity price shocks generate significant U-shaped responses in output, consumption, and per-capita hours. Most notably, the inflation rate displays a significant spike, followed by a rapid return to the initial level. The sudden surge in the inflation rate prompts the Fed to elevate the nominal interest rate. Results of a counterfactual exercise (in the style of Bernanke et al., 1997) indicate that the systematic contractionary response helped the Federal Reserve to achieve price stability in the long run, yet at the cost of a significant economic downturn in output and per-capita hours.

Our framework also addresses the hours response to neutral technology shocks. In particular, we find that the response of per-capita hours is positive and marginally significant. This result is surprising, given that we control for low-frequency movements in the data (Canova et al., 2010). Further investigations show that this result is very robust to the treatment of the data as long as the size of the information set is sufficiently large. This result, which is in line with the evidence found by Forni and Gambetti (2011), confirms our choice to estimate a large-scale SVAR.

The sub-sample properties of our model are consistent with Blanchard and Galí (2010). We find that the effects of a commodity price shock on output and the inflation rate are milder in the post-Volcker period, but remain statistically significant at the 10% level. Several further robustness checks confirm the findings of our model. In particular, we examine robustness to the choice of the lag length, the identifying assumptions, the specific commodity price index used, and the inclusion of an external demand shock (Abbritti and Weber, 2010).

Appendix 2.A Additional Tables

Table 5: Sources and Definitions of Data

Series	Definition	Source	Mnemonic
POP	civilian non-institutional population 16+	FRED	CNP16OV
FFR	effective (net) Federal Funds rate	FRED	FEDFUNDS
CPI	consumer price index (all urban consumers)	FRED	CPIAUCSL
PPI	producer price index (crude materials)	FRED	PPICRM
GOV	real government consumption expenditures & gross investment	FRED	GCEC96
EXP	real exports of goods & services	FRED	EXPGSC1
IMP	real imports of goods & services	FRED	IMPGSC1
HOU	hours in the business sector	BLS	PRS84006033
OUT	real output per hour in the business sector	BLS	PRS84006093
EMP	employment in the business sector	BLS	PRS84006013
RPI	quality-adjusted relative price of investment	DiCecio (2009)	p_i
CON	real personal consumption expenditures (nondurables & services)	DiCecio (2009)	cndq + csq
INV	real quality adjusted gross private fixed investment + PCE durables, divided by 100	DiCecio (2009)	r_inv
CCI	continuous commodity index	Datastream	NYFECRB

Notes: This table displays the definitions of the raw series used. The BLS (2012, p. 5) defines crude materials for further processing as "[. . .] unprocessed commodities not sold directly to consumers. Crude foodstuffs and feedstuffs include items such as grains and livestock. The crude energy goods category consists of crude petroleum, natural gas to pipelines, and coal. Examples of crude nonfood materials other than energy include raw cotton, construction sand and gravel, and iron and steel scrap". Current and historical weights can be downloaded at: `ftp : //ftp.bls.gov/pub/special.requests/ppi/`, e.g. `sopnew08.txt` summarizes the weights in December 2007. We also thank Riccardo DiCecio for kindly sharing his data. The quality-adjustment follows Gordon (1990), Cummins and Violante (2002), and Fisher (2006). Consumer durables are included in investment, but the change in inventories is not. We aggregate all monthly series to quarterly data.

Table 6: Definition of Variables in the SVAR

Variable	Symbol	Definition
growth in labor p'tivity	Δa_t	first difference of log (OUT)
growth in RPI	Δq_t	first difference of log (RPI)
per-capita hours	h_t	log of (HOU/POP)
inflation rate	π_t	first difference of log (CPI)
nominal interest rate	r_t	FFR
employment rate	n_t	log of (EMP/POP)
commodity price index	p_t	log of (PPI)
consumption share	c_t	log of (CON/(CON+INV+GOV+EXP-IMP))
investment share	i_t	log of (INV/(CON+INV+GOV+EXP-IMP))
export/import ratio	d_t	log of (EXP/IMP)
continuous commodity index	cci_t	log of (CCI)

Notes: This table displays the variables that enter the SVAR. The trivariate model (Canova et al., 2010) uses only the first three variables. The last two variables are only used for robustness checks.

Bibliography

Abbritti, Mirko and Sebastian Weber, "Labor market institutions and the business cycle: Unemployment rigidities vs. real wage rigidities," Working Paper Series No. 1183, European Central Bank 2010.

Abel, Andrew B., "Asset Prices under Habit Formation and Catching Up with the Joneses," *American Economic Review*, 1990, *80* (2), 38–42.

Alquist, Ron, Lutz Kilian, and Robert J. Vigfusson, "Forecasting the Price of Oil," in Graham Elliott and Allan Timmermann, eds., *Handbook of Economic Forecasting*, 2 ed., Vol. 24 of *Handbooks in Economics*, Elsevier, 2011. Forthcoming.

Altig, David, Lawrence J. Christiano, Martin Eichenbaum, and Jesper Lindé, "Firm-Specific Capital, Nominal Rigidities and the Business Cycle," *Review of Economic Dynamics*, 2011, *14* (2), 225–247.

Anzuini, Alessio, Marco J. Lombardi, and Patrizio Pagano, "The impact of monetary policy shocks on commodity prices," Working Paper Series No. 1232, European Central Bank 2010.

Barro, Robert J. and Robert G. King, "Time-separable preferences and intertemporal-substitution models of business cycles," *Quarterly Journal of Economics*, 1984, *99* (4), 817–839.

Barsky, Robert B. and Lutz Kilian, "Do we Really Know that Oil Caused the Great Stagflation? A Monetary Alternative," in Ben S. Bernanke and K. Rogoff, eds., *NBER Macroeconomics Annual 2001*, Vol. 16 of *NBER Macroeconomics Annual*, Cambridge, MA: The MIT Press, 2002, pp. 137–183.

Basu, Susanto, "Intermediate Goods and Business Cycles: Implications for Productivity and Welfare," *American Economic Review*, 1995, *85* (3), 512–531.

Baumeister, Christiane and Gert Peersman, "The Role of Time-Varying Price Elasticities in Accounting for Volatility Changes in the Crude Oil Market," *Journal of Applied Econometrics*, 2012. Forthcoming.

_ **and** _ , "Time-Varying Effects of Oil Supply Shocks on the U.S. Economy," Working Papers No. 12-2, Bank of Canada 2012.

Bernanke, Ben S., Mark Gertler, and Mark W. Watson, "Systematic Monetary Policy and the Effects of Oil Price Shocks," *Brookings Papers on Economic Activity*, 1997, *1*, 91–142.

Bils, Mark and Peter J. Klenow, "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 2004, *112* (5), 947–985.

Blanchard, Olivier J. and Danny Quah, "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 1989, *79* (4), 655–673.

_ **and Jordi Galí**, "The Macroeconomic Effects of Oil Price Shocks: Why are the 2000s so Different from the 1970s?," in Jordi Galí and Mark Gertler, eds., *International Dimensions of Monetary Policy 2007*, Chicago, IL: University of Chicago Press, 2010, pp. 373–421.

Blinder, Alan S. and Jeremy B. Rudd, "The Supply-Shock Explanation of the Great Stagflation Revisited," NBER Working Papers No. 14563, National Bureau of Economic Research 2008.

BLS, *BLS Handbook of Methods*, Bureau of Labor Statistics, U.S. Department of Labor, 2012. Chapter 14: Producer Prices.

Canova, Fabio, David Lopez-Salido, and Claudio Michelacci, "The effects of technology shocks on hours and output: A robustness analysis," *Journal of Applied Econometrics*, 2010, *25* (5), 755–773.

Christiano, Lawrence J. and Terry J. Fitzgerald, "The Band Pass Filter," *International Economic Review*, 2003, *44* (2), 435–465.

_ , Martin Eichenbaum, and Charles L. Evans, "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," *Review of Economics and Statistics*, 1996, *78* (1), 16–34.

_ , _ , and _ , "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 2005, *113* (1), 1–45.

_ , _ , and Robert J. Vigfusson, "What Happens After a Technology Shock?," NBER Working Paper Series 9819, National Bureau of Economic Research, Cambridge, MA July 2003.

_ , _ , and _ , "The Response of Hours to a Technology Shock: Evidence Based on Direct Measures of Technology," *Journal of the European Economic Association*, 2004, *2* (2-3), 381–395.

Cummins, Jason and Gianluca Violante, "Investment-Specific Technical Change in the United States (1947-2000)," *Review of Economic Dynamics*, 2002, *5* (2), 243–248.

De Bock, Reinout, "Investment-Specific Technology Shocks and Labor Market Frictions," Research series No. 200701-01, National Bank of Belgium January 2007.

Dedola, Luca and Stefano Neri, "What does a technology shock do? A VAR analysis with model-based sign restrictions," *Journal of Monetary Economics*, 2007, *54* (2), 512–549.

di Pace, Federico and Renato Faccini, "Deep habits and the cyclical behaviour of equilibrium unemployment and vacancies," *Journal of Economic Dynamics and Control*, 2011. Forthcoming.

DiCecio, Riccardo, "Sticky wages and sectoral labor comovement," *Journal of Economic Dynamics and Control*, 2009, *33* (3), 538–553.

Dupor, Bill, Jing Han, and Yi-Chan Tsai, "What do technology shocks tell us about the New Keynesian paradigm?," *Journal of Monetary Economics*, 2009, *56* (4), 560–569.

Eichenbaum, Martin, "Comments on "Interpreting the macroeconomic time series facts: The effects of monetary policy" by Christopher A. Sims," *European Economic Review*, 1992, *36* (5), 1001–1011.

Enders, Zeno, Gernot J. Müller, and Almuth Scholl, "How do fiscal and technology shocks affect real exchange rates? New evidence for the United States," *Journal of International Economics*, 2011, *83* (1), 53–69.

Erceg, Christopher J., Luca Guerrieri, and Christopher Gust, "Can long-run restrictions identify technology shocks?," *Journal of the European Economic Association*, 2005, *3* (6), 1237–1278.

Evans, Charles L. and Jonas D. M. Fisher, "What are the implications of rising commodity prices for inflation and monetary policy?," Chicago Fed Letter No. 286, The Federal Reserve Bank of Chicago 2011.

Faccini, Renato and Salvador Ortigueira, "Labor-Market Volatility in the Search-and-Matching Model: The Role of Investment-Specific Technology Shocks," *Journal of Economic Dynamics and Control*, 2010, *34* (8), 1509–1527.

Fernald, John G., "Trend breaks, long-run restrictions, and contractionary technology improvements," *Journal of Monetary Economics*, 2007, *54* (8), 2467–2485.

Fisher, Jonas D. M., "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks," *Journal of Political Economy*, 2006, *114* (3), 413–451.

Forni, Mario and Luca Gambetti, "Testing for Sufficient Information in Structural VARs," CEPR Discussion Paper No. 8209, Centre for Economic Policy Research, London 2011.

Francis, Neville and Valerie A. Ramey, "Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited," *Journal of Monetary Economics*, November 2005, *52* (8), 1379–1399.

_ and _ , "Measures of per Capita Hours and Their Implications for the Technology-Hours Debate," *Journal of Money, Credit and Banking*, 2009, *41* (6), 1071–1097.

Furlanetto, Francesco and Martin Seneca, "Investment-specific technology shocks and consumption," Working Paper No. 2010/30, Norges Bank 2010.

Galí, Jordi, "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review*, 1999, *89* (1), 249–271.

_ and Paul Rabanal, "Technology Shocks and aggregate fluctuations: How well does the RBC model fit postwar U.S. data?," in Mark Gertler and Kenneth Rogoff, eds., *NBER Macroeconomics Annual 2004*, Vol. 19 of *NBER Macroeconomics Annual*, Cambridge, MA: The MIT Press, 2005, pp. 225–288.

Gordon, Robert J., *The Measurement of Durable Goods Prices* NBER Monograph, Chicago, IL: University of Chicago Press, 1990.

Gospodinov, Nikolay, Alex Maynard, and Elena Pesavento, "Sensitivity of Impulse Responses to Small Low Frequency Co-Movements: Reconciling the Evidence on the Effects of Technology Shocks," *Journal of Business and Economic Statistics*, 2011, *29* (4), 455–467.

Granger, Clive W. J., "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 1969, *37* (3), 424–438.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell, "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, 1997, *87* (3), 342–362.

_ , _ , and _ , "The role of investment-specific technological change in the business cycle," *European Economic Review*, 2000, *44* (1), 91–115.

Hamilton, James D., *Time Series Analysis*, Princeton, New Jersey: Princeton University Press, 1994.

_ , "Historical Oil Shocks," NBER Working Paper Series 16790, National Bureau of Economic Research, Cambridge, MA February 2011.

_ **and Ana Maria Herrera**, "Oil Shocks and Aggregate Economic Behavior: The Role of Monetary Policy: Comment," *Journal of Money, Credit and Banking*, 2004, *36* (2), 265–286.

Harris, Ethan S., Bruce C. Kasman, Matthew D. Shapiro, and Kenneth D. West, "Oil and the Macroeconomy: Lessons for Monetary Policy," in "Conference of the US Monetary Policy Forum (USMPF)" University of Chicago (Booth) and Brandeis University (IBS) 2009.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti, "Investment shocks and business cycles," *Journal of Monetary Economics*, 2010, *57* (2), 132–145.

Kilian, Lutz, "Confidence intervals for impulse responses under departures from normality," *Econometric Reviews*, 1998a, *17* (1), 1–29.

_ , "Small-Sample Confidence Intervals For Impulse Response Functions," *Review of Economics and Statistics*, 1998b, *80* (2), 218–230.

_ , "The economic effects of energy price shocks," *Journal of Economic Literature*, 2008, *46* (4), 871–909.

_ , "Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market,," *American Economic Review*, 2009, *99* (3), 1053–1069.

_ **and Logan Lewis**, "Does the Fed Respond to Oil Price Shocks?," *Economic Journal*, 2011, *121* (555), 1047–1072.

_ **and Robert J. Vigfusson**, "Are the Responses of the U.S. Economy Asymmetric in Energy Price Increases and Decreases?," *Quantitative Economics*, 2011, *2* (3), 419–453.

King, Robert G. and Sergio T. Rebelo, "Resuscitating real business cycles," in J. B. Taylor and M. Woodford, eds., *Handbook of Macroeconomics*, 1a ed., Vol. 15 of *Handbooks in Economics*, Elsevier, April 1999, chapter 14, pp. 927–1007.

Kydland, Finn E. and Edward C. Prescott, "Time to Build and Aggregate Fluctuations," *Econometrica*, 1982, *50* (6), 1345–1370.

Lucas, Robert E., "Econometric policy evaluation: A critique," *Carnegie-Rochester Conference Series on Public Policy*, 1976, *1* (1), 19–46.

_ , "Two Illustrations of the Quantity Theory of Money," *American Economic Review*, 1980, *70* (5), 1005–1014.

Lütkepohl, Helmuth, *New Introduction to Multiple Time Series Analysis*, Berlin: Springer, 2005.

Mackowiak, Bartosz and Mirko Wiederholt, "Business Cycle Dynamics under Rational Inattention," Working Paper Series No. 1331, European Central Bank 2011.

Mumtaz, Haroon and Francesco Zanetti, "Neutral technology shocks and the dynamics of labor input: Results from an agnostic identification," *International Economic Review*, 2011. Forthcoming.

Ordóñez, Javier, Hector Sala, and Jose I. Silva, "Oil price shocks and labor market fluctuations," *The Energy Journal*, 2011, *32* (3), 89–118.

Ravn, Morten O. and Saverio Simonelli, "Labor Market Dynamics and the Business Cycle: Structural Evidence for the United States," *Scandinavian Journal of Economics*, 2008, *109* (4), 743–777.

_ , Stephanie Schmitt-Grohé, Martín Uribe, and Lenno Uusküla, "Deep habits and the dynamic effects of monetary policy shocks," *Journal of The Japanese and International Economies*, 2010, *24* (2), 236—-258.

Rotemberg, Julio J., "Cyclical wages in a search-and-bargaining model with large firms," in Lucrezia Reichlin and Kenneth West, eds., *NBER International Seminar on Macroeconomics 2006*, Chicago, IL: University of Chicago Press, 2008, pp. 65–114.

_ and Michael Woodford, "Imperfect competition and the effects of energy price increases on economic activity," *Journal of Money, Credit and Banking*, 1996, *28* (4), 459–577.

Sargan, John D., "Wages and prices in the United Kingdom: A Study in Econometric Methodology," in P. E. Hart, G. Mills, and J. K. Whitaker, eds., *Econometric Analysis for National Economic Planning*, Vol. 16 of *Colston Papers*, London: Butterworth, 1964, pp. 25–63.

Schmitt-Grohé, Stephanie and Martín Uribe, "Business Cycles With A Common Trend in Neutral and Investment-Specific Productivity," *Review of Economic Dynamics*, 2011, *14* (1), 122–135.

Shapiro, Matthew D. and Mark W. Watson, "Sources of Business Cycles Fluctuations," in Stanley Fisher, ed., *NBER Macroeconomics Annual 1988*, Vol. 3 of *NBER Macroeconomics Annual*, Cambridge, MA: The MIT Press, 1988, pp. 111–156.

Sims, Christopher A., "Macroeconomics and Reality," *Econometrica*, 1980, *48* (1), 1–48.

_ , "Interpreting the macroeconomic time series facts: The effects of monetary policy," *European Economic Review*, 1992, *36* (5), 975–1000.

Smets, Frank and Rafael Wouters, "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 2007, *97* (3), 586–606.

Thomson Reuters, "Thomson Reuters Equal Weight Continuous Commodity Index (CCI)," 2010. **URL:** `http://thomsonreuters.com/content/financial/pdf/i_and_a/indices/cci_factsheet.pdf`.

Thomson Reuters/Jefferies, "Thomson Reuters/Jefferies CRB Index," 2011. **URL:** `http://thomsonreuters.com/content/financial/pdf/i_and_a/indices/trj_crb.pdf`.

Uhlig, Harald, "Do Technology Shocks Lead to a Fall in Total Hours Worked?," *Journal of the European Economic Association*, 2004, *2* (2-3), 361–371.

Chapter 3

The Balassa-Samuelson Effect Reversed: New Evidence from OECD Countries

Abstract

This chapter explores the robustness of the Balassa-Samuelson (BS) hypothesis. We analyze a panel of OECD countries from 1970 to 2008 and compare three different data sets on sectoral productivity, including a newly constructed database on total factor productivity. Overall, our DOLS estimation results do not support the BS hypothesis. For the last two decades, we find a very robust negative equilibrium relationship between the productivity in the tradable sector and the real exchange rate, in contrast to BS. Earlier supportive findings depend strongly on the choice of the data set. Except for the terms of trade, the explanatory power of other variables is weak.

Keywords: Real Exchange Rate, Balassa-Samuelson Hypothesis, Panel Data Estimation, Terms of Trade

1 Introduction

The Balassa-Samuelson (BS) hypothesis is one of the most widespread explanations for structural deviations from purchasing power parity (PPP) (Dornbusch, 1985).[1] The BS hypothesis was stated by both Balassa (1964) and Samuelson (1964) simultaneously, but has a research precedent in the work of Harrod (1933). According to the hypothesis, differences in the productivity differential between the non-tradable and the tradable sector lead to differences in price levels between countries, when converted to the same currency.[2] Ceteris paribus, a productivity increase in tradables raises factor prices, i.e., wages, which in turn leads to higher prices of non-tradables and thus to an appreciation of the real exchange rate. In contrast, when the relative productivity of non-tradables increases, marginal cost cuts result in a lower price level.

The empirical evaluation of the BS hypothesis has gained a great deal of attention. As argued in a survey by Tica and Družić (2006), the major share of the evidence supports the BS model, but the strength of the results depends on the nature of the tests and set of countries analyzed.

There are several studies based on a disaggregation of the tradable and non-tradable sector that find empirical support for the BS hypothesis (see, e.g., Calderón, 2004; Choudhri and Khan, 2005 or Lee et al., 2008). In particular, since sector-specific data for OECD countries on total factor productivity (TFP) has become available, various studies have tested and confirmed the BS hypothesis using OECD panel data (De Gregorio and Wolf, 1994; Chinn and Johnston, 1996; MacDonald and Ricci, 2007). All these studies are based on the discontinued International Sectoral Database (ISDB) provided by the OECD. According to this literature, countries with higher productivity differentials between tradables and non-tradables exhibit higher price levels expressed in a single currency, i.e., a stronger real exchange rate.

This chapter estimates the long-run relationship between the real exchange rate and key explanatory variables, including the TFP differential between tradables and non-tradables. We apply a panel cointegration model that manages to treat the non-stationarity of the variables correctly. Recently, the OECD has provided a new database called PDBi with sector-specific TFP data from 1985 to 2008.

[1] According to the absolute PPP theory, a unit of currency must have the same purchasing power in the foreign economy as in the domestic economy, once it is converted into the same currency. However, in his seminal work establishing the so-called PPP puzzle, Rogoff (1996) argues that the speed of adjustment of real exchange rates is too slow to be in line with the PPP theory. Recent studies that stress the importance of nonlinear adjustments (Taylor, 2003) or dynamic aggregation bias (Imbs et al., 2005) challenge this finding. Indeed, there is empirical evidence for the failure of PPP, although the results are mixed (for reviews see Froot and Rogoff, 1996 or Taylor, 2003).

[2] The hypothesis assumes that the law of one price for tradable goods holds. However, there are empirical contributions to the literature that find deviations from this law (see, e.g., Engel, 1999) as well as theoretical explanations for these deviations (for a discussion see MacDonald and Ricci (2007) and the references therein).

With this new data set, our estimations cannot confirm the findings of previous research based on the ISDB.[3] In fact, the results point to a negative relationship between tradable productivity and the real exchange rate. This finding is the *opposite* of what is claimed by the BS hypothesis. Furthermore, we can confirm this result when TFP is replaced by labor productivity (LP) using the OECD Structural Analysis (STAN) database, which covers more countries and a longer time period, from 1970 to 2008. On the other hand, the connection between non-tradable productivity and the real exchange rate is not robust. Finally, with the exception of the terms of trade, our estimation results indicate that the explanatory power of control variables discussed in the literature is weak or not robust.

In order to detect the causes of the conflicting results, we systematically compare the three data sets and their implications regarding the estimation results. Our robustness tests reveal that severe outlier dependency exists for the traditional Balassa-Samuelson finding regarding *non-tradables*. In particular, Japanese labor productivity in the non-tradable sector strongly weakens the estimated BS effect. For the time period from 1970 to 1992, the coefficient even significantly changes its sign once Japan is included, a finding that is robust against several variations of the model specification. For the time period after 1992, the connection strongly depends on the country sample, particularly wether the United States are included or not.

However, the negative relationship between the productivity of tradables and the real exchange rate is very robust. A rigorous analysis of the tradable sector reveals that this reversal is robust for the last two decades against the choice of the country sample, the precise start of the sample period, the exact model specification, and the inclusion of additional explanatory variables. In Gubler and Sax (2012), we provide a static general-equilibrium framework with skill-based technological change (SBTC) that can explain the new relationship. We show that SBTC may increase the relative amount of physical capital used by the tradable industry, thus releasing low-skilled labor; and finally, leading to lower wages for low-skilled labor and lower prices in the economy. An increase in tradable productivity may thus be connected to a lower real exchange rate.

Overall, we conclude from the results of our analysis that the presence of the Balassa-Samuelson effect is not robust for a panel of major OECD countries. In fact, during the last two decades an increase in the productivity of tradables, all else being equal, has given rise to a depreciation of the real exchange rate.

The remainder of this chapter is organized as follows. Section 2 presents the data. We outline the methodology in Section 3 and show the results in Section 4. Section 5 concludes.

[3]However, we are indeed able to replicate the results in favor of the BS hypothesis with data from the ISDB.

2 Data

The data for the 18 major OECD countries included in our data set stem from different databases of the IMF, OECD, World Bank and the Penn World Tables. Depending on the estimation, the country sample has to be reduced, because we aim to replicate the results of MacDonald and Ricci (2007) or because not all data are available.[4] A detailed description of all variables is given in Table (1) and in Appendix 3.A.2.

Table 1: Description and Construction of the Variables

Abbr.	Name	Definition	Source
RER	Real Exchange Rate	log(CPI / Nominal Exchange Rate to USD)	IMF, IFS
TFP_T	TFP of Tradables	Solow Residual	OECD, EO
TFP_NT	TFP of Non-Tradables	Solow Residual	OECD, EO
LP_T	LP of Tradables	log(Value Added / Hours-Worked)	OECD, EO
LP_NT	LP of Non-Tradables	log(Value Added / Hours-Worked)	OECD, EO
CA	Current Account	as % of GDP	OECD, EO
DPOP	Population Growth	Δ log(Population)	PWT
GDP	Real GDP per capita	log(Real GDP per capita)	PWT
GOV	Government Spending	as % of GDP	OECD, EO
NFA	Net Foreign Assets	as % of GDP	WB, WDI
RI	Long-Term Real Int. Rate	Gov. bond yield long term - CPI	IMF, IFS
TOT	Terms of Trade	log(Export-Prices / Import-Prices)	OECD, EO

In order to test the BS hypothesis, we condition the real exchange rate on productivity measures for both the tradable and the non-tradable sector, as well as on control variables. The choice of the dependent variable is discussed in Section 2.1. Due to its importance and complexity, productivity data are separately examined in Section 2.2. All other exogenous variables are discussed in Section 2.3. The time series properties of the variables are assessed in the final Section 2.4.

2.1 Dependent Variable: Real Exchange Rate

We use the logarithm of the unweighted real exchange rate, *RER*, as the dependent variable in our estimation equations. In principle, the real exchange rate can only be computed towards a reference country. However, instead of defining such a reference country, we circumvent the problem by using time fixed effects throughout our analysis. The real exchange rate can thus be thought of as a deviation from the sample mean, which, in this case, acts as the reference country ("average OECD country"). Proceeding this way allows us to keep as many observations as possible.

An extensive body of the empirical literature uses *effective* real exchange rates (see, e.g., De Gregorio and Wolf, 1994; Calderón, 2004 or Lee et al., 2008), that are weighted by the share of exports. Effective real exchange rates have the advantage that there is

[4] All country samples featured in our estimations are presented in Appendix 3.A.1.

no need to specify a specific reference country. While effective real exchange rates are a useful measure for competitiveness, the share of exports seems not only irrelevant in our context but also misleading. If, for example, a country changes its export destinations to countries with a weaker real exchange rate, effective real exchange rates would indicate a real appreciation, while, in fact, the country still has the same relative price level towards all countries.

2.2 Productivity Data

We use data on sectoral productivity from three data sets provided by the OECD: The first is a new data set on sectoral total factor productivity (TFP) computed by the OECD, called PDBi (Productivity Database, i represents the specific sector). The data set contains annual sector-specific TFP numbers and covers the time period from 1985 to 2008. Sectoral TFP is calculated as Solow residuals by the same method for all countries, using sectoral data on production, employment, capital stock and the labor share of income. Capital stocks are estimated applying the permanent inventory method, where streams of investments are added, and a certain fraction of depreciation is subtracted each year.

A second database, STAN, includes yearly data on sectoral production and employment, and thus on labor productivity, but not on capital stock or TFP. As the only data set, STAN covers a long time range from 1970 to 2008 for many OECD countries.

In order to compare our findings with existing studies, in particular with the results of MacDonald and Ricci (2007), sectoral productivity data from the discontinued International Sectoral Database (ISDB) have been used as well. This old database contains annual values on labor and total factor productivity, in principle from 1970 to 1997, but has been discontinued before 1997 for most countries.

STAN and PDBi data are improvements to the ISDB. In the old data set, output, employment and capital stocks were based on data from an old system of national accounts, SNA68. For social services, these changes in the measurement of output may have been especially important, as estimates of real value added growth for the public sector in the ISDB have simply been based on labor inputs, so that estimates of productivity had a very limited meaning. Moreover, in the ISDB, volumes were calculated using constant prices instead of chained linking. Finally, capital stock estimates may have been calculated differently and in a non-standardized way in the ISDB.[5]

The classification of subsectors into tradable and non-tradable is made according to the following scheme: Agriculture, manufacturing, and transport, storage and communications are classified as tradables; utilities (energy, gas and water), construction, and social services (community, social, personal services) as non-tradables. Our division of

[5] While these are very general observations about the evolution of the system of national accounts, it would be desirable if international organizations could provide more information about the changes over time. As in the present case, that would tremendously facilitate the task of replicating earlier results.

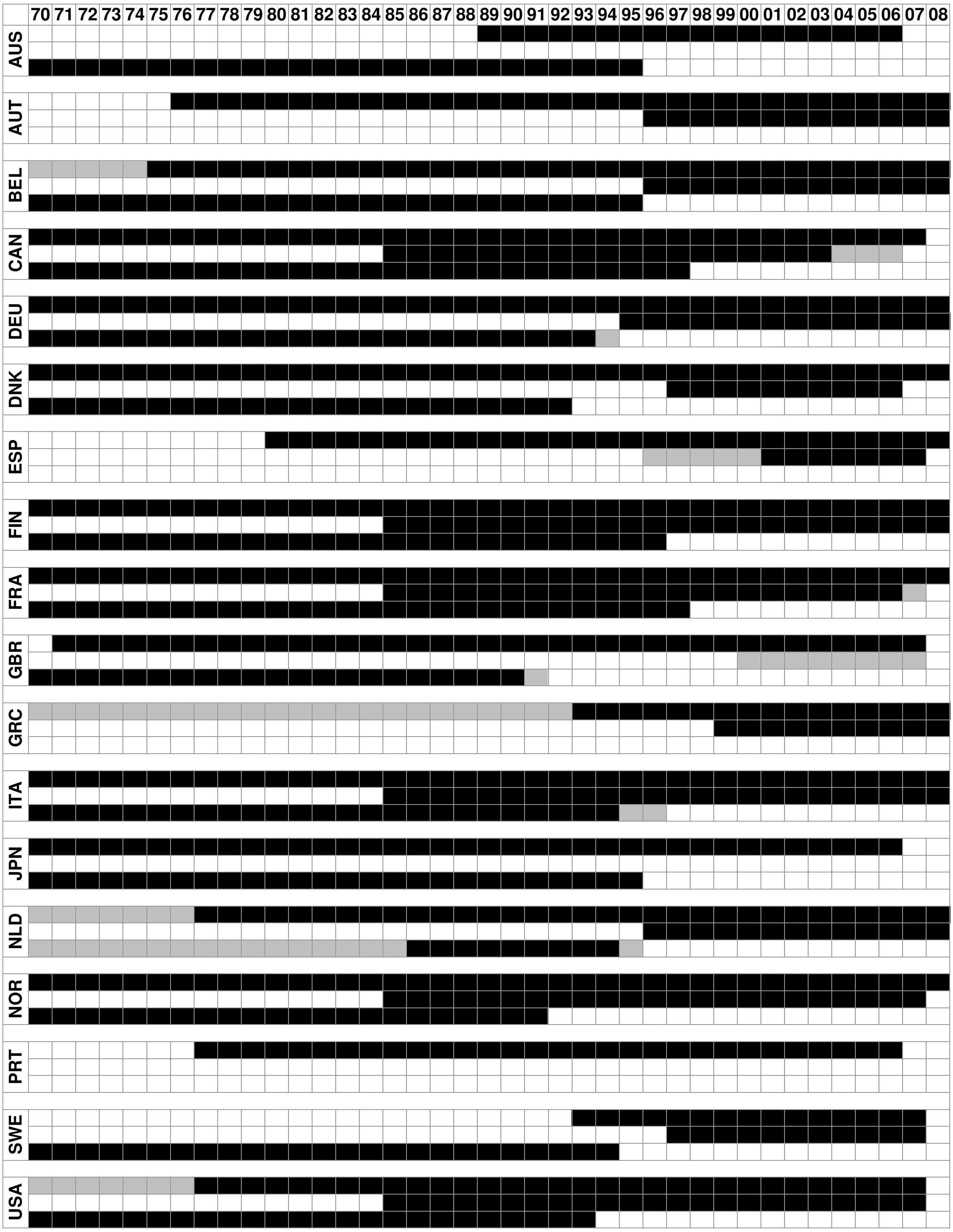


Figure 1: Sectoral Productivity Data Coverage

Notes: For each country, the first row describes the coverage span of the STAN database, the second of the PDBi and the third of the ISDB. If all six sectors are available, the line is drawn *black*, if some sectors are available, it is drawn *grey*. The STAN database covers the broadest range of the three databases.

Table 2: Median Correlations across Subsectors

	AGR	IND	TSC	EGW	CST	SOC
PDBi (TFP), STAN (LP)	0.95	0.97	0.92	0.95	0.93	0.84
ISDB (TFP), STAN (LP)	0.90	0.91	0.93	0.75	0.76	**0.28**
ISDB (TFP), ISDB (LP)	0.99	0.98	0.98	0.96	0.94	0.97
ISDB (LP), STAN (LP)	0.90	0.88	0.88	0.72	0.77	**0.27**
ISDB (EMP), STAN (EMP)	0.91	0.98	0.91	0.89	0.99	**0.45**
ISDB (VA), STAN (VA)	0.91	0.95	0.89	0.72	0.93	**0.45**

Notes: The table contains median correlation coefficients between variables in the three data sets for all six subsectors. The values are based on all countries for which a correlation coefficient can be calculated. AGR: Agriculture; IND: Manufacturing; TSC: Transport, Storage and Communications; EGW: Energy, Gas and Water; CST: Construction; SOC: Community, Social, Personal Services. The first three rows show the median correlations between TFP from the PDBi or the ISDB and LP from the STAN database or the ISDB. Median correlations between LP from the STAN database and the ISDB are reported in the fourth row. The last two rows contain the median correlation values between EMP from the ISBD and the STAN database, and between VA from the same sources. Due to a very low number of time-overlapping observations, no comparison between the PDBi and the ISDB is presented.

the subsectors into tradable or non-tradable sectors follows De Gregorio and Wolf (1994), who defined a subsector as tradable if its share of exports in the total production exceeds 10% and as non-tradable otherwise.[6] While no division has become standard in the field (Tica and Družić, 2006), studies based on data from OECD countries usually refer to the division proposed by De Gregorio and Wolf (1994) (see, e.g., Chinn and Johnston, 1996; MacDonald and Ricci, 2007). Like MacDonald and Ricci (2007), we exclude distribution, mining and financial subsectors due to classification difficulties (MacDonald and Ricci, 2005), or data availability.

The tradable and non-tradable sectors, classified this way, are roughly equal in terms of value added. Each sector comprises 50% of the total value added produced by these six subsectors. Within the tradable sector, manufacturing is by far the largest subsector, representing 64% of value added, whereas both agriculture and transport, storage and communications amount to 11% and 24%, respectively. Among the non-tradables, social services (70%) outweigh construction (20%) and utilities (9%). Figure (1) displays the data availability in each of the three data sets.

Table (2) shows the correlations between the three data sets. LP and TFP values from the ISDB are similar to the two newer data sets only in the tradable subsectors. In the non-tradable sectors, the correlations are lower (construction and utilities) or virtually non-existent (social services). To a lesser extent, this is also true for employment and value added. Possible reasons for these divergences have been discussed earlier in this section. On the other hand, data from the PDBi on TFP are highly correlated with labor productivity from the STAN database. These correlations are present in all subsectors, although the values are somewhat lower in the non-tradable subsectors.

[6] Adjustments of the threshold value to 5% and 20% leave the division virtually unchanged (De Gregorio and Wolf, 1994).

We consider TFP as the preferred measure for productivity. As pointed out by De Gregorio and Wolf (1994), average labor productivity grows much faster during economic downturns, and hence, it is not a reliable indicator of sustainable productivity growth which can affect the economy in the medium or long term. Nevertheless, there are some advantages of LP, and we will use the measure to check the robustness of our TFP results.[7]

2.3 Control Variables

We include several control variables along with the productivity variables in our estimations: The terms of trade, TOT, has been proposed to be an important determinant of the long-run real exchange rate (see, e.g., De Gregorio and Wolf, 1994; Sax and Weder, 2009). An improvement in the terms of trade allows a country to raise its imports for a given amount of factor inputs in the export sector. Hence, a change in TOT may be interpreted analogous to a change in the productivity in the tradable sector (De Gregorio and Wolf, 1994).

Several authors point out the importance of demand-side factors for the determination of the long-run real exchange rate. Therefore, we consider the government spending share, GOV, net foreign assets, NFA, the current account, CA, and real GDP per capita, GDP, as control variables.

De Gregorio and Wolf (1994) show theoretically that an increase in government spending reduces the equilibrium real exchange rate if capital mobility across countries is restricted. This increase affects the relative price of tradable and non-tradable goods negatively, because government spending tends to fall more heavily on non-tradables. Hence, government spending is widely used as an additional explanatory variable (see, e.g., Chinn and Johnston, 1996; Lee et al., 2008 or Sax and Weder, 2009).

Private demand may affect the real exchange rate as well. It is likely that a higher income is associated with a higher demand for non-tradables. The associated rise in the relative price of non-tradables gives rise to a higher overall price level (De Gregorio and Wolf, 1994). Furthermore, trade deficits or surpluses could affect the demand for non-tradables, by increasing or decreasing the amount of tradables that are available for consumption. As a permanent trade deficit can only be sustained in the presence of net foreign assets, several authors have emphasized the importance either of the net foreign assets or the current account deficit for the determination of the real exchange rate (Krugman, 1990; Lane and Milesi-Ferretti, 2004; Lee et al., 2008).

Finally, two other macroeconomic variables, the real interest rate, RI, and the population growth rate, $DPOP$, are taken into account. Their importance for the determination

[7]The advantages are summarized by Canzoneri et al. (1999): First, labor productivity data are available for more countries and over a longer time period than TFP numbers. Second, the calculation of LP figures does not require an estimation of the capital stock and the income share of labor, with both estimations likely to be imprecise. Third, the BS hypothesis holds for more technologies than the Cobb-Douglas production function generally employed to determine TFP.

Table 3: IPS and LLC Panel Unit Root Test Results

		Det. Trend	IPS	LLC	No. of Countries	Time Period	Obs.
CA			0.933	0.994	18	1970-2008	587
$DPOP$			-4.269***	-2.837***	18	1970-2007	626
GDP		x	1.010	1.591	18	1970-2007	656
GOV		x	3.091	0.130	18	1970-2008	632
NFA			3.920	5.589	18	1970-2006	611
RER_AVG		x	-1.172	-1.116	18	1970-2008	665
RI			-0.500	-0.331	18	1970-2008	621
TOT			0.233	0.214	18	1970-2008	640
LP_T	STAN	x	1.282	-1.540*	18	1970-2008	559
LP_NT	STAN	x	1.651	1.131	18	1970-2008	550
TFP_T	PDBi	x	-0.021	-1.537*	14	1985-2008	198
TFP_NT	PDBi	x	1.782	0.077	13	1985-2008	192
LP_T	ISDB	x	2.923	2.906	14	1970-1997	325
LP_NT	ISDB	x	1.909	1.103	14	1970-1997	322
TFP_T	ISDB	x	1.360	0.886	14	1970-1997	314
TFP_NT	ISDB	x	1.720	0.614	14	1970-1997	307

Notes: x indicates the inclusion of a deterministic trend. As all estimations contain time-specific dummy variables, the real exchange rate of each country is computed with respect to the average sample country for the unit root tests (RER_AVG). IPS: Lag length selection by modified SIC (Ng and Perron, 2001); LLC: Lag length selection by modified SIC; Bartlett kernel, Newey-West bandwidth. The panel is unbalanced: The time period marks the maximum years available */**/*** denotes significance at 10%, 5% and 1% level, respectively

of RER has been discussed in theoretical and empirical contributions to the literature. According to the theoretical model provided by Stein and Allen (1997), a higher real interest rate is associated with an appreciated long-run real exchange rate because of portfolio adjustments and capital inflows. Rose et al. (2009) show in an overlapping generation model that a country experiencing a decline in its fertility rate will also experience a real exchange rate depreciation.

2.4 Assessing the Time Series Properties of the Variables

The panel unit root tests proposed by Levin et al. (2002) (LLC) and Im et al. (2003) (IPS) have been conducted for all variables (Table 3). In order to obtain reliable results, the test statistics are based on all available information for both time and cross-sectional dimensions. Thus, although the full length of a series is not used for the estimation, all observations are nevertheless used for the unit root tests.

As described in Section 2.1, we do not compute the real exchange rate towards a specific country (or towards a basket of countries, such as in the effective real exchange rate). Instead, time-specific dummy variables are included in all estimations. In order to assess the times series properties, the real exchange rate is calculated towards the average of the sample, denoted RER_AVG.

Overall, we find strong evidence for non-stationary behavior for all variables with the exception of the population growth rate, $DPOP$. As $DPOP$ is the first difference of the logarithm of population, this result is not surprising. The total factor productivity in the tradable sector, TFP_T, from the PDBi and labor productivity in the tradable sector, LP_T, from the STAN database show ambiguous results. However, non-stationarity of these variables is confirmed by the Fisher-type augmented Dickey-Fuller (ADF) panel unit root test proposed by Maddala and Wu (1999) (not shown) and is theoretically founded in macroeconomic models (see, e.g., King et al., 1991; Galí, 1999 or Lindé, 2009). All results are also in line with the results found in similar empirical studies (see, e.g., Calderón, 2004; MacDonald and Ricci, 2007 or Lee et al., 2008).

3 Methodology

The number of observations for each country is limited, given the length of the sample (23 years in our benchmark model) and the annual data frequency. Therefore, we pool the data and apply a panel estimation technique to improve the power of our results. We are primarily interested in the long-run relationship between the real exchange rate and its determinants described in Section 2 and summarized in Table (1).

In order to estimate this relationship, we employ a panel cointegration model that treats the non-stationarity of the variables correctly. Furthermore, the dynamic speed of adjustment of the real exchange rate to its long-run equilibrium is determined.

Our estimation results are based on the dynamic ordinary least squares (DOLS) method. Several methods to estimate a panel cointegration model are discussed in the literature. However, Kao and Chiang (2000) show that the DOLS approach developed by Stock and Watson (1993) outperforms the panel OLS or the fully modified OLS (FMOLS) procedures in the sense that the DOLS estimator is less biased in finite samples. In addition, the choice of this method facilitates comparison with the results from similar studies, e.g., MacDonald and Ricci (2007). Our estimation equation has the following form:

$$RER_{it} = \delta_t + \alpha_i + X_{it}\beta + \sum_{j=-p}^{j=k} \Delta X_{it+j}\gamma_j + \epsilon_{it} \tag{1}$$

where RER_{it} denotes the real exchange rate at time t of country i, α_i is a country fixed effect, δ_t is a time fixed effect, X_{it} is a vector containing the explanatory variables, β is the cointegration vector, k and p are the maximum and minimum lag lengths, respectively, γ_j are the $k+p+1$ vectors containing the coefficients of the leads and lags of changes in the explanatory variables, and ϵ_{it} represents the error term.

The inclusion of the leads and lags solves the potential endogeneity problem by orthogonalizing the error term.[8] We choose the number of leads and lags to be one ($k = p = 1$).

[8]The leads and lags remove the correlation between the error term and the stationary component of

A rising number of leads and lags further constrains the number of observations. This may be a caveat particularly in subsamples with reduced numbers of years.[9]

Both time and country fixed effects are added in order to reduce omitted variable bias, and because some variables are indices, and hence, their levels are not comparable across countries. Furthermore, as described in Section 2.1, time fixed effects are necessary, because our real exchange rate is not computed towards a reference country.

We report standard errors developed by Driscoll and Kraay (1998) that are robust to very general forms of spatial and temporal dependence. For the computation, we follow Cribari-Neto (2004), who proposed an estimator (called HC4) that is reliable when the data contain influential observations.[10]

To ensure that what we find is indeed a long-run relationship between the real exchange rate and the set of explanatory variables, we test for cointegration using two methods. First, we follow MacDonald and Ricci (2007), who apply the unit root test of Levin et al. (2002) to the estimated residuals.[11] Second, we employ the Kao (1999) panel cointegration test. Since this test requires a balanced panel, some observations have to be dropped, and, therefore, the test is mainly applied to check the robustness of the first test results.

In order to capture the short-run dynamic adjustment of the real exchange rate to temporary disequilibria, an error correction specification is applied to the data. The model can be written as follows:

$$\begin{aligned}\Delta RER_{it} &= \Theta_i + \Theta_t + \eta\, gap_{it-1} + \\ &\quad \sum_{j=0}^{j=1} \phi_j \Delta RER_{it-j} + \sum_{j=0}^{j=1} \Delta X_{it-j}\omega_j + \mu_{it}\end{aligned} \tag{2}$$

where

$$gap_{it} = RER_{it} - \delta_t - \alpha_i - X_{it}\beta \tag{3}$$

and where η represents the adjustment speed coefficient. The gap_{it} is computed using the results of Equation (1).

4 Empirical Results

In order to explore the validity of the Balassa-Samuelson (BS) hypothesis, we estimate various DOLS model specifications. This section presents the results for the long-run

the non-stationary variables.

[9] However, our main conclusions are robust to an increased number of leads and lags.

[10] As a robustness check, we employ the HC3 estimator proposed by Long and Ervin (2000). The conclusions do not change.

[11] For the theoretical foundation of this methodology, see Pedroni (2004). The conclusions do not change if the residuals are corrected by the estimated leads and lags.

relationship between the real exchange rate and relative productivity as well as control variables. In addition, the results of the cointegration tests described in Section 3 are reported. Finally, we provide an extensive robustness analysis of our main findings.

4.1 Comparison of the Productivity Data Sets

In a first step, we examine the validity of the BS hypothesis using the newest sectoral productivity data from the PDBi. For the purpose of comparability, we restrict our sample to the same set of countries and control variables as MacDonald and Ricci (2007). Therefore, the real exchange rate, RER, is conditioned on total factor productivity or labor productivity of tradables (TFP_T, LP_T) and non-tradables (TFP_NT, LP_NT), net foreign assets in percent of GDP, NFA, and the long-term real interest rate, RI. The cross-section dimension is reduced to the countries listed in sample (i) in Appendix 3.A.1.

The results in column (1) of Table (4) are based on the model with TFP data from the new Productivity Database (PDBi) and the sample period lasting from 1985 to 2008.[12] There is a statistically significant negative impact of TFP_T on RER. A 10% increase in the TFP of tradables implies a 4% depreciation of the real exchange rate. The coefficient on TFP_NT is positive but not significant. Overall, this result contradicts not only the BS hypothesis, but also the usual conclusions drawn from empirical studies analyzing Balassa-Samuelson in a panel of OECD countries.

However, most of the related literature obtains sectoral productivity data from the older International Sectoral Database (ISDB). For a comparison with the existing literature, in particular MacDonald and Ricci (2007), column (2) reports the estimation results with TFP data from the ISDB for the period from 1970 to 1992. Except for NFA, the results are now qualitatively equal to the findings of MacDonald and Ricci (2007). In particular, the signs of the coefficients related to both TFP variables are consistent with the BS hypothesis. Quantitatively, the coefficients on both coefficients are smaller than in MacDonald and Ricci (2007), and the coefficient on TFP_T is statistically insignificant. These differences disappear to a great extent once we follow MacDonald and Ricci (2007) and increase the number of leads and lags to three. The results are shown in column (1) of Table (8) in Appendix 3.B. Hence, we are able to replicate the results in favor of the BS theory with data from the ISDB.

The successful confirmation of the BS hypothesis may depend either on the productivity data source or the sample period or both. Unfortunately, the two data sets ISDB and PDBi contain only very few overlapping observations. Therefore, we are not able to distinguish the *time* from the *source* effect. In order to verify our finding, we estimate the model with labor productivity (LP) data from STAN, as this database covers both periods.[13]

[12]Notice that Japan is not covered by the PDBi database.

[13]Notice that due to lack of data for some years, the coverage is not exactly the same. See Figure (1)

Table 4: Comparison of the Data Sets (DOLS)

Dependent Variable: *RER*					
	(1)	(2)	(3)	(4)	(5)
TFP_T	-0.427***	0.119			
	(0.108)	(0.280)			
TFP_NT	0.260	-0.661***			
	(0.227)	(0.101)			
LP_T			-0.114***	-0.114	-0.196***
			(0.039)	(0.161)	(0.058)
LP_NT			0.580***	1.058***	0.445
			(0.290)	(0.170)	(0.383)
NFA	0.0004	-0.002	0.0002	-0.001	0.0001
	(0.001)	(0.002)	(0.001)	(0.004)	(0.001)
RI	0.006	0.007	0.008	0.014***	-0.007
	(0.009)	(0.007)	(0.010)	(0.004)	(0.012)
LLC Test	-6.385***	-6.632***	-6.522***	-5.388***	-5.934***
Kao Test	-3.647***	-4.043***	-3.030***	-2.588***	-3.663***
η	-0.46	-0.24	-0.18	-0.21	-0.47
Half Lifetime (years)	1.1	2.5	3.4	2.9	1.1
Sample Period	1985-2008	1970-1992	1970-2008	1970-1992	1992-2008
Obs.	112	188	289	163	113

Notes: See Table (1) for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country sample (Appendix 3.A.1): Sample (i). Japan is not included in (1) (see Footnote 12). The productivity data stem from the PDBi (1), the ISDB (2), and the STAN database (3)-(5). Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. LLC test: Cointegration test following MacDonald and Ricci (2007): t-statistic of Levin et al. (2002) (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). Kao test: Cointegration test proposed by Kao (1999): t-statistic (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). η is obtained from Equation (2). Half lifetime of deviations of the real exchange rate from estimated relation (years): ln(0.5)/(ln(1 + η)). */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Column (3) displays the results for the whole sample from 1970 to 2008. Column (4) shows the results of the subsample from 1970 to 1992 and thus covers the same period as the estimations with productivity data from the ISDB (column 2). The second subsample (column 5) ranges from 1992 to 2008.

The coefficients on LP_T are negative in all estimations and statistically significant for the whole sample and for the second subsample, confirming the results from column (1). The coefficients on LP_NT are positive in all estimations and statistically significant for the whole sample and for the first subsample, again confirming the results from column (1). Thus, labor productivity data from the STAN database lead to similar results as total factor productivity data from the PDBi, but both results contradict the BS hypothesis and differ from the findings using the ISDB. This result is further analyzed in Section 4.4. As the sign of the coefficients on LP_T and LP_NT is the same across both subsamples, the differences between columns (1) and (2) are likely to be determined by the data set rather than by the sample period. Note that the essential difference remains when we

for more details.

substitute LP for TFP from the ISDB. The coefficients are displayed in column (2) of Table (8) in Appendix 3.B.

The control variable NFA has the correct sign in columns (1), (3) and (5). However, the coefficients are statistically insignificant and the economic effect is considerably smaller compared to the results of Lane and Milesi-Ferretti (2004) and Lee et al. (2008). The model with labor productivity data from the STAN database estimated from 1970 to 1992 has a significant positive coefficient on RI consistent with the theory (Stein and Allen, 1997).

Our results with ISDB productivity data point to a half-life of deviations of the real exchange rate from its estimated long-run relationship of 2.5 to 2.7 years (column 2, Table 4, and column 2, Table 8, in Appendix 3.B). This result is in line with the existing literature (see, e.g., Lee et al., 2008), but larger than in MacDonald and Ricci (2007). In recent times, the adjustment speed has accelerated to about one year. These results are reported in columns (1) and (5) of Table (4). As our main interest concerns the cointegration relationship, we do not discuss this issue further.

4.2 Full Country Sample Estimations

The estimations in Table (4) are based on a reduced set of countries. We now re-estimate the model with all countries available (sample (ii) and (iii), Appendix 3.A.1). In addition, we drop the variables NFA and RI, since neither variable seems to have considerable explanatory power for the long-run real exchange rate. Instead, we use the terms of trade, TOT, as a control variable in the baseline model, as TOT turns out to be an important and robust determinant of the real exchange rate.[14]

Table (5) summarizes the results. Compared to Table (4), the coefficients on TFP_T and LP_T are smaller, but remain negative and significant with a single insignificant exception (column 3). However, for the same first period subsample, the coefficient is also insignificant (but negative) with the reduced country sample (column 4, Table 4). A 10% increase in the TFP of tradables from the new PDBi would imply an almost 2% depreciation of the real exchange rate. A similar result emerges with LP data from the STAN database for the sample period from 1992 to 2008 (column 4). Thus, the negative relationship between the productivity of tradables and the real exchange rate persists when all countries are included. We will further explore the robustness of this relationship in Section 4.4.

The effect of non-tradable productivity is less robust. Compared to Table (4), the coefficients on the productivity of non-tradables switch signs in two estimations: In col-

[14]The inclusion of TOT raises the concern about possible endogeneity. We conduct a very simple exercise to check for reverse causation by substituting the contemporaneous value by the one-year-lagged value of TOT. The results are robust to this modification. Therefore, we conclude that this potential endogeneity problem is not of a major concern in our analysis. The results are shown in Table (9) in Appendix 3.B.

Table 5: Full Country Sample Estimation Results (DOLS)

Dependent Variable: *RER*	(1)	(2)	(3)	(4)
TFP_T	-0.176***			
	(0.061)			
TFP_NT	-0.767**			
	(0.314)			
LP_T		-0.0493**	0.143	-0.164***
		(0.024)	(0.143)	(0.062)
LP_NT		0.596***	0.407***	-0.110
		(0.185)	(0.125)	(0.151)
TOT	0.265*	0.302**	0.132	0.263**
	(0.145)	(0.126)	(0.156)	(0.127)
LLC Test	-8.160***	-8.766***	-8.468***	-9.002***
Kao Test	-2.052**	1.285*	1.352*	1.622*
Sample Period	1985-2008	1970-2008	1970-1992	1992-2008
Obs.	181	532	251	258

Notes: See Table (1) for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country samples (Appendix 3.A.1): Sample (ii) for (1), sample (iii) for (2) and (4), and sample (iv) for (3). The productivity data stem from the PDBi (1), and the STAN database (2)-(4). Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. LLC test: Cointegration test following MacDonald and Ricci (2007): t-statistic of Levin et al. (2002) (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). Kao test: Cointegration test proposed by Kao (1999): t-statistic (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). */**/*** denotes significance at 10%, 5% and 1% level, respectively.

umn (1), the coefficient on TFP_NT drops from 0.26 to -0.77 and becomes statistically significant (compared to column 1, Table 4), in column (4) the coefficient on LP_NT turns negative, but remains insignificant (compared to column 5, Table 4). We will explore this *lack* of robustness in Section 4.4.

TOT is statistically and economically significant with the correct sign in columns (1), (2) and (4). On average, a 10% increase in the terms of trade leads to a 2% appreciation of the real exchange rate.

4.3 Effects of Control Variables

The impact of additional explanatory variables on the long-run real exchange rate is analyzed in Table (6). In line with the previous results, both coefficients on the productivity variables are negative and predominantly significant in all models. For the tradable sector productivity, this is the opposite effect of what is claimed by the BS hypothesis. Additionally, the significant positive impact of the terms of trade on the price level remains.

The selection of the explanatory variables is discussed in Section 2.3. In line with the theory, government spending, GOV, has a positive but insignificant effect on RER (column 1). Moreover, a current account surplus, CA, has a statistically significant positive effect, as predicted (column 2); however, the very small coefficient points to a limited

Table 6: Control Variables (DOLS)

Dependent Variable: *RER*				
Variables	(1)	(2)	(3)	(4)
TFP_T	-0.140	-0.083	-0.211**	-0.239***
	(0.087)	(0.090)	(0.084)	(0.060)
TFP_NT	-0.696**	-0.336	-0.701**	-0.845***
	(0.317)	(0.281)	(0.355)	(0.286)
TOT	0.231**	0.396**	0.177*	0.254
	(0.109)	(0.170)	(0.106)	(0.163)
GOV	0.001			
	(0.002)			
CA		-0.010***		
		(0.003)		
GDP			0.254***	
			(0.089)	
DPOP				-15.8
				(9.93)
LLC Test	-8.316***	-8.150***	-7.166***	-7.871***
Kao Test	-1.817**	-2.595***	-3.102***	-2.521***
Obs.	181	181	174	169

Notes: See Table (1) for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country sample (Appendix 3.A.1): Sample (ii). The productivity data stem from the PDBi. Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. LLC test: Cointegration test following MacDonald and Ricci (2007): t-statistic of Levin et al. (2002) (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). Kao test: Cointegration test proposed by Kao (1999): t-statistic (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). */**/*** denotes significance at 10%, 5% and 1% level, respectively.

economic significance. Real GDP per capita, GDP, affects RER significantly positive (column 3) and confirms the hypothesis that the income level affects the consumption pattern: A 10% increase in GDP implies a 3% increase appreciation of the real exchange rate. Finally, contrary to the theory, in our sample of OECD countries, there is no significant connection between the population growth rate, $DPOP$, and RER (column 4).

4.4 Robustness Analysis

From the estimation results presented in the previous sections, we conclude that an overall stable Balassa-Samuelson effect cannot be found. The coefficients on productivity of tradables and, in particular, on the productivity of non-tradables are not robust against various sample variations, such as the source of the productivity data, the time period, or the set of control variables.

As we have seen, TFP from the PDBi in the non-tradable sector changes its coefficient from 0.26 to -0.77 and becomes highly significant once we employ the full country sample (column 1, Table 4, and column 1, Table 5). By repeatedly re-estimating this specification and each time omitting one of the countries of sample (ii) (Appendix 3.A.1), we are able

Table 7: The Impact of Japan on the Results (DOLS)

Dependent Variable: *RER*			
Variables	(1)	(2)	(3)
LP_T	-0.070**	-0.020	-0.189***
	(0.027)	(0.113)	(0.051)
LP_NT	0.304***	-0.334**	-0.256*
	(0.112)	(0.179)	(0.133)
TOT	0.380***	0.397***	0.165**
	(0.075)	(0.090)	(0.075)
LLC Test	-9.230***	-8.561***	-9.832***
Kao Test	0.350	2.223**	1.125
Obs.	497	230	245

Notes: See Table (1) for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country samples (Appendix 3.A.1): Sample (iii) without Japan for (1) and (3), sample (iv) without Japan for (2). The productivity data stem the STAN database. Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. LLC test: Cointegration test following MacDonald and Ricci (2007): t-statistic of Levin et al. (2002) (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). Kao test: Cointegration test proposed by Kao (1999): t-statistic (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). */**/*** denotes significance at 10%, 5% and 1% level, respectively.

to identify the United States as a critical outlier. As soon as the country is omitted, the coefficient changes from being significantly negative to being insignificantly positive. But if we continue the exclusion exercise without the United States, we find that omitting Italy switches the sign of the coefficient again, this time from positive to negative. Then, excluding France changes the coefficient to positive. Next, dropping Norway leads to another sign reversal. Although the coefficients are insignificant, at least for the non-tradable sector, the result crucially depends on the country sample chosen.

The estimation of individual slope coefficients on non-tradable productivity confirms this finding. While the effect is positive for Austria, Denmark, Greece, Italy and Norway, the contrary holds for Belgium, Finland, France, Germany, Netherlands, Spain, Sweden and the United States. Again, only the coefficient on the United States is significant. Therefore, the relationship between non-tradable productivity and the real exchange rate seem to differ across the countries making a panel approach for analyzing this relationship questionable.

Furthermore, Japan seems to be an outlier that critically affects the results of the estimations using labor productivity data from the STAN database. Re-estimating the specifications in columns (2)-(4) of Table (5) without Japan produces different results, reported in Table (7). In particular, while an increase in labor productivity in the non-tradable sector gives rise to a significant real exchange rate appreciation in the full country sample from 1970 to 1992 (column 3, Table 5), the contrary is true if Japan is omitted (column 2, Table 7).

The replication of the exclusion exercise with country sample (iv) in the absence of

Japan shows that the coefficient on LP_NT remains negative and becomes insignificant only when Norway or the United States are excluded. Therefore, the positive relationship between Japanese non-tradable productivity and RER dominates the overall negative relationship in the other countries.

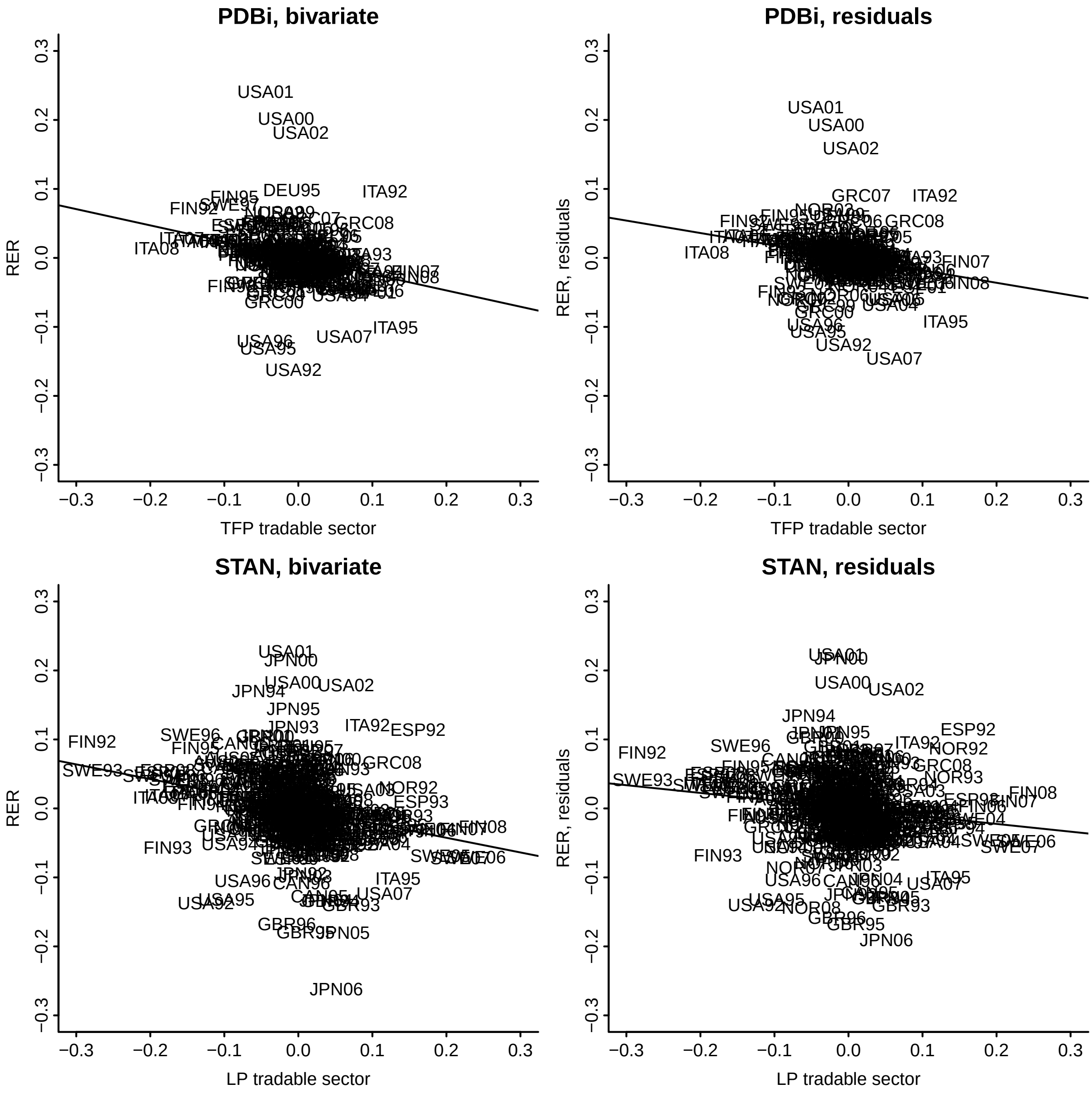


Figure 2: Tradable Productivity and the Real Exchange Rate since 1992

This result is also robust against moving the end date from 1985 to 1995, although the coefficient shrinks and loses its significance for the years 1993 until 1995. Furthermore, the sign remains unchanged with additional explanatory variables taken into account (Table 10 in Appendix 3.B). Finally, the increase in the number of leads and lags of up to three (MacDonald and Ricci, 2007) does not alter the conclusion.

There is a second, truly robust finding: The relationship between the productivity in the tradable sector and the real exchange rate is negative, since the beginning of the 1990s (columns 1 and 4, Table 5). While this result contradicts the BS hypothesis, it is robust against variations in the country sample, the sample period, the exact model specification and the set of control variables. The finding is also independent of whether we use labor or total factor productivity.

The exclusion exercise reveals that the negative sign is persistent against the omission of any country. In rare cases, the coefficient becomes statistically insignificant (if Italy, Sweden (with LP), or Norway (with LP and TFP) are excluded). Varying the start point of the sample shows that the negative coefficient is significantly negative from 1988 to 1995, independent of the productivity data source. The relationship remains significantly negative when up to three leads and lags are included into the model. Finally, the finding is robust against the inclusion of additional explanatory variables (Tables 11 and 12 in Appendix 3.B).

Once we increase the starting point of the sample from 1985 to 1992, the variables CA and GDP are not significant anymore (Table 12 in Appendix 3.B). Of all additional explanatory variables, it is thus only the terms of trade that is robust against a sample variation.

The relationship between the productivity in the tradable sector and the real exchange rate is illustrated in Figure (2). The left panel contains the productivity of tradables in relation to the real exchange rate, both adjusted by country-specific and time-specific effects. In the right panel, the real exchange rate is additionally adjusted by the productivity of non-tradables and the terms of trade. The small differences between the left and the right panel indicate that the relationship does not depend on whether control variables are used. In line with the DOLS estimation results, all scatter plots show the significant negative relationship.

Having found a robust negative relationship between tradable productivity and the real exchange rate, we conclude that sectoral productivity panel data from the OECD offer no support for the BS hypothesis for a panel of major OECD countries.

5 Summary and Conclusions

This chapter explores the robustness of the Balassa-Samuelson (BS) hypothesis. We analyze a panel of OECD countries from 1970 to 2008 and compare three different data sets on sectoral productivity provided by the OECD, including a newly constructed database on total factor productivity (TFP).

Overall, we cannot find support for the BS hypothesis. In contrast, our DOLS estimations point to a very robust negative equilibrium relationship between the productivity in the tradable sector and the real exchange rate for the last two decades. We find this negative relationship with respect to TFP from the new Productivity Database (PDBi) as well as with sectoral labor productivity (LP) from the STAN database. The finding not only contradicts the BS hypothesis, but also the results of previous empirical research that is based on the older International Sectoral Database (ISDB).

Results from estimations with LP indicate that the difference in the findings from studies using TFP data from the ISDB (in favor of BS) and the PDBi (against BS) are due to the data source and not due to a change of the relationship over time.

An extensive robustness analysis shows that the negative relationship does not depend on the choice of the country sample, the precise start of the time period, the exact model specification, the inclusion of additional explanatory variables or the non-tradable productivity. On the other hand, the relationship between the productivity in the non-tradable sector and the long-run real exchange rate during the last two decades is strongly affected by the choice of the country sample.

Prior to 1992, the robustness tests further reveal a strong dependency of the results on a single outlier: The coefficient on non-tradable labor productivity significantly changes the sign once Japan is included. Without Japan, we find a robust negative relationship between non-tradable productivity and the real exchange rate, in line with the BS hypothesis.

Finally, we examine the explanatory power of control variables whose importance for the real exchange rate determination has been discussed in the literature. The results indicate that, with the exception of the terms of trade, their explanatory power is weak or not robust against the chosen time period.

The fact that we find a robust negative relationship between tradable productivity and the real exchange rate is puzzling. According to the Balassa-Samuelson hypothesis, we would expect a higher productivity to be connected with higher wages and thus with a higher price level. Why is the opposite the case? In Gubler and Sax (2012), we show that skill-based technological change may lead to an increase in productivity *and* a lower demand for low-skilled labor, and thus to lower prices in the economy. Of course, other explanations are equally possible.

Our findings potentially facilitate future empirical research on the determination of the equilibrium real exchange rate in OECD countries. As the Balassa-Samuelson hypothesis does not contribute to an explanation of real exchange rate movements, sectoral productivity data do not have to serve necessarily as a control variable. This should bring a major gain in data availability. Not only more countries but more years can be included without running into potential omitted variable bias.

Appendix 3.A Data Appendix

3.A.1 Country Samples

This section contains all country samples used in the estimation models:

i Belgium (BEL), Denmark (DNK), Finland (FIN), France (FRA), Germany (DEU), Italy (ITA), Japan (JPN), Norway (NOR) and Sweden (SWE)

ii Austria (AUT), Belgium (BEL), Denmark (DNK), Finland (FIN), France (FRA), Germany (DEU), Greece (GRC), Italy (ITA), Netherlands (NLD), Norway (NOR), Spain (ESP), Sweden (SWE) and the United States (USA)

iii Australia (AUS), Austria (AUT), Belgium (BEL), Canada (CAN), Denmark (DNK), Finland (FIN), France (FRA), Germany (DEU), Great Britain (GBR), Greece (GRC), Italy (ITA), Japan (JPN), Netherlands (NLD), Norway (NOR), Portugal (PRT), Spain (ESP), Sweden (SWE) and the United States (USA)

iv Austria (AUT), Belgium (BEL), Canada (CAN), Denmark (DNK), Finland (FIN), France (FRA), Great Britain (GBR), Greece (GRC), Italy (ITA), Japan (JPN), Netherlands (NLD), Norway (NOR), Portugal (PRT), Spain (ESP) and the United States (USA)

3.A.2 Data Sources

i IMF, International Financial Statistics

We gained the following IFS variables via Datastream:

- BOND YIELD (AUY61... etc.)
- CPI (AUY64...F etc.)
- EXCHANGE RATE, US$ PER LC (AUOCFEXR etc.)

ii OECD, Economic Outlook

The data are from Economic Outlook No 88., available on `http://www.oecd-ilibrary.org/`. These variables were used:

- Imports of goods and services, deflator, national accounts basis (PMGSD)
- Exports of goods and services, deflator, national accounts basis (PXGSD)
- Current account balance, as a percentage of GDP (CBGDPR)
- Total disbursements, general government, as a percentage of GDP

iii OECD, STAN Database for Structural Analysis

The data are from `oecd.stat` and have been downloaded as a single ASCII file. The series for Germany have been retropolated with the former West-Germany series.

iv OECD, PDBi, Sectoral Productivity Database

A new data set provided by the OECD (we used a pre-released version of the data set).

Both for the STAN database and the PDBi, tradable and non-tradable productivity is calculated the following way:

$$\begin{aligned} P_{NT} &= \frac{S_{7599} \cdot P_{7599} + S_{4041} \cdot P_{4041} + S_{4500} \cdot P_{4500}}{S_{7599} + S_{4041} + S_{4500}}, \\ P_{T} &= \frac{S_{0105} \cdot P_{0105} + S_{1537} \cdot P_{1537} + S_{6064} \cdot P_{6064}}{S_{0105} + S_{1537} + S_{6064}}, \end{aligned}$$

where P denotes labor productivity in the STAN case and total factor productivity in the PDBi case. S is the share of the subsector.

v OECD, ISDB, Sectoral Productivity Database

A vintage data set provided by the OECD.

Tradable and non-tradable total factor productivity is calculated the following way (again, P denotes labor or total factor productivity and S the share of the subsector):

$$\begin{aligned} P_{NT} &= \frac{S_{SOC} \cdot P_{SOC} + S_{EGW} \cdot P_{EGW} + P_{CST} \cdot S_{CST}}{S_{SOC} + S_{EGW} + S_{CST}}, \\ P_{T} &= \frac{S_{AGR} \cdot P_{AGR} + S_{MAN} \cdot P_{MAN} + S_{TRS} \cdot P_{TRS}}{S_{AGR} + S_{MAN} + S_{TRS}}. \end{aligned}$$

vi Penn World Tables

The data are from `http://pwt.econ.upenn.edu/php_site/pwt_index.php`. These variables were used:

- Real GDP-per-capita (USD of 2005) (RGDPL)
- Population (in 1000) (POP)

The population growth rate is calculated as the first difference of the logarithm of POP.

vii World Bank, World Development Indicators

The following variables are extracted from the WDI CD-ROM:

- Net foreign assets

The share of net foreign assets (NFA in the text) is calculated in the following way:

$$NFA = \frac{NFA_{Level}}{GDP \cdot 1000000}$$

where NFA_{Level} are the net foreign assets as taken from WDI and GDP denotes nominal GDP taken from the OECD Economic Outlook. The missing value of NFA_{Level} for Belgium and France for the year 1998 is replaced by a linearly interpolated value. The results do not change.

Appendix 3.B Additional Tables

Table 8: Additional Results with ISDB Productivity Data (DOLS)

Dependent Variable: RER	(1)	(2)
TFP_T	0.958***	
	(0.328)	
TFP_NT	-1.070***	
	(0.122)	
LP_T		0.430**
		(0.182)
LP_NT		-0.166**
		(0.074)
NFA	0.0004	0.004*
	(0.003)	(0.002)
RI	-0.010	0.009
	(0.010)	(0.008)
LLC Test	-6.146***	-6.632***
Kao Test	-4.043***	-1.285*
η	-0.23	-0.23
Half Lifetime (years)	2.7	2.7
Sample Period	1970-1992	1970-1992
Obs.	152	188

Notes: See Table (1) for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable. (1) includes 3 leads/lags; (2) includes 1 lead/lag. Country sample (Appendix 3.A.1): Sample (i). The productivity data stem from the ISDB. Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. LLC test: Cointegration test following MacDonald and Ricci (2007): t-statistic of Levin et al. (2002) (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). Kao test: Cointegration test proposed by Kao (1999): t-statistic (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). η is obtained from Equation (2). Half lifetime of deviations of the real exchange rate from estimated relation (years): $\ln(0.5)/(\ln(1+\eta))$. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Table 9: Full Country Sample Estimation Results with $TOT(-1)$ (DOLS)

Dependent Variable: RER	(1)	(2)	(3)	(4)
TFP_T	-0.147**			
	(0.070)			
TFP_NT	-0.893***			
	(0.296)			
LP_T		-0.046*	0.140	-0.219***
		(0.024)	(0.142)	(0.049)
LP_NT		0.527***	0.374**	-0.282*
		(0.194)	(0.180)	(0.146)
$TOT(-1)$	0.285	0.377***	0.210	0.271**
	(0.175)	(0.121)	(0.150)	(0.132)
Sample Period	1985-2008	1970-2008	1970-1992	1992-2008
Obs.	168	514	237	240

Notes: See Table (1) for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country samples (Appendix 3.A.1): Sample (ii) for (1) and sample (iii) for (2)-(4); Australia, Germany and Sweden are not included in (3) due to missing data. The productivity data stem from the PDBi (1), and the STAN database (2)-(4). Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Table 10: Control Variables: Japan Omitted (DOLS)

Dependent Variable: RER Variables	(1)	(2)	(3)	(4)
LP_T	-0.053	-0.179	-0.089	-0.046
	(0.107)	(0.141)	(0.127)	(0.153)
LP_NT	-0.177	-0.107	-0.257**	-0.329**
	(0.125)	(0.188)	(0.118)	(0.138)
TOT	0.412***	0.480***	0.390***	0.427***
	(0.122)	(0.044)	(0.086)	(0.085)
GOV	0.001			
	(0.002)			
CA		-0.0138***		
		(0.004)		
GDP			0.726***	
			(0.136)	
$DPOP$				-2.71
				(2.31)
Obs.	229	195	230	224

Notes: See Table (1) for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country sample (Appendix 3.A.1): Sample (iv) without Japan. The productivity data stem from the STAN database. Sample period: 1970-1992. Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Table 11: Control Variables: Estimations with LP (DOLS)

Dependent Variable: *RER*				
Variables	(1)	(2)	(3)	(4)
LP_T	-0.142**	-0.083	-0.184***	-0.209***
	(0.064)	(0.079)	(0.071)	(0.045)
LP_NT	-0.083	-0.071	-0.050	-0.206
	(0.149)	(0.150)	(0.156)	(0.162)
TOT	0.249**	0.321***	0.258***	0.319***
	(0.113)	(0.118)	(0.098)	(0.089)
GOV	0.002			
	(0.002)			
CA		-0.006***		
		(0.002)		
GDP			0.296**	
			(0.120)	
DPOP				-7.78
				(8.8)
Obs.	258	258	247	231

Notes: See Table (1) for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country sample (Appendix 3.A.1): Sample (iii). The productivity data stem from the STAN database. Sample period: 1992-2008. Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Table 12: Control Variables: Estimations with TFP (DOLS)

Dependent Variable: *RER*				
Variables	(1)	(2)	(3)	(4)
TFP_T	-0.222***	-0.221**	-0.377***	-0.284***
	(0.083)	(0.089)	(0.067)	(0.070)
TFP_NT	-0.912**	-0.886**	-0.954*	-0.655
	(0.433)	(0.408)	(0.530)	(0.405)
TOT	0.168***	0.195**	0.131*	0.036
	(0.060)	(0.089)	(0.068)	(0.059)
GOV	0.001			
	(0.003)			
CA		-0.002		
		(0.002)		
GDP			0.256	
			(0.167)	
DPOP				-22.6*
				(13.7)
Obs.	146	146	139	134

Notes: See Table (1) for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country sample (Appendix 3.A.1): Sample (ii). The productivity data stem from the PDBi. Sample period: 1992-2008. Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Bibliography

Balassa, Bela, "The Purchasing-Power Parity Doctrine: A Reappraisal," *The Journal of Political Economy*, December 1964, *72* (6), 584–596.

Calderón, César A., "Real Exchange Rates in the Long and Short Run: A Panel Cointegration Approach," *Revista de Análisis Económico*, 2004, *19*, 41–83.

Canzoneri, Matthew B., Robert E. Cumby, and Behzad Diba, "Relative labor productivity and the real exchange rate in the long run: evidence for a panel of OECD countries," *Journal of International Economics*, 1999, *47* (2), 245–266.

Chinn, Menzie D. and Louis Johnston, "Real Exchange Rate Levels, Productivity and Demand Shocks: Evidence from a Panel of 14 Countries," NBER Working Paper 5709, National Bureau of Economic Research, Cambridge, MA August 1996.

Choudhri, Ehsan U. and Mohsin S. Khan, "Real Exchange Rates in Developing Countries: Are Balassa-Samuelson Effects Present?," *IMF Staff Papers*, 2005, *52* (3), 387–409.

Cribari-Neto, Francisco, "Asymptotic inference under heteroskedasticity of unknown form," *Computational Statistics & Data Analysis*, 2004, *45* (2), 215–233.

De Gregorio, José and Holger C. Wolf, "Terms of Trade, Productivity, and the Real Exchange Rate," NBER Working Paper 4807, National Bureau of Economic Research, Cambridge, MA July 1994.

Dornbusch, Rudiger, "Purchasing Power Parity," NBER Working Paper 1591, National Bureau of Economic Research, Cambridge, MA March 1985.

Driscoll, John C. and Aart C. Kraay, "Consistent covariance matrix estimation with spatially dependent panel data," *Review of Economics and Statistics*, 1998, *80* (4), 549–560.

Engel, Charles, "Accounting for US Real Exchange Rate Changes," *Journal of Political Economy*, 1999, *107* (3), 507–538.

Froot, Kenneth A. and Kenneth Rogoff, "Perspectives on PPP and Long-Run Real Exchange Rates," NBER Working Paper 4952, National Bureau of Economic Research, Cambridge, MA April 1996.

Galí, Jordi, "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review*, 1999, *89* (1), 249–271.

Gubler, Matthias and Christoph Sax, "Skill-Biased Technological Change and the Real Exchange Rate," Unpublished Manuscript, University of Basel, 2012.

Harrod, Roy F., *International Economics* Cambridge Economic Handbooks, London: Nisbet & Cambridge University Press, 1933.

Im, Kyung S., M. Hashem Pesaran, and Yongcheol Shin, "Testing for unit roots in heterogeneous panels," *Journal of Econometrics*, 2003, *115* (1), 53–74.

Imbs, Jean, Haroon Mumtaz, Morten O. Ravn, and Hélène Rey, "PPP Strikes Back: Aggregation and the Real Exchange Rate," *Quarterly Journal of Economics*, 2005, *120* (1), 1–43.

Kao, Chihwa, "Spurious regression and residual-based tests for cointegration in panel data," *Journal of Econometrics*, 1999, *90* (1), 1–44.

_ **and M. H. Chiang**, "Nonstationary panels, cointegration in panels and dynamic panels: A survey," *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, 2000, *15*, 179–222.

King, Robert G., Charles I. Plosser, and James H. Stock, "Stochastic Trends and Economic Fluctuations," *The American Economic Review*, 1991, *81* (4), 819–840.

Krugman, Paul, "Equilibrium Exchange Rates," in William H. Branson, Jacob A. Frenkel, and Morris Goldstein, eds., *International Policy Coordination and Exchange Rate Fluctuations*, The University of Chicago Press, 1990.

Lane, Philip R. and Gian Maria Milesi-Ferretti, "The transfer problem revisited: Net foreign assets and real exchange rates," *Review of Economics and Statistics*, 2004, *86* (4), 841–857.

Lee, Jaewoo, Gian Maria Milesi-Ferretti, and Luca Antonio Ricci, "Real Exchange Rates and Fundamentals: A Cross-Country Perspective," *IMF Working Papers*, 2008.

Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu, "Unit root tests in panel data: asymptotic and finite-sample properties," *Journal of Econometrics*, 2002, *108* (1), 1–24.

Lindé, Jesper, "The effects of permanent technology shocks on hours: Can the RBC-model fit the VAR evidence?," *Journal of Economic Dynamics and Control*, March 2009, *33* (3), 597–613.

Long, J. Scott and Laurie H. Ervin, "Using heteroscedasticity consistent standard errors in the linear regression model," *American Statistician*, 2000, pp. 217–224.

MacDonald, Ronald and Luca Antonio Ricci, "The Real Exchange Rate and the Balassa–Samuelson Effect: The Role of the Distribution Sector," *Pacific Economic Review*, 2005, *10* (1), 29–48.

_ **and** _ , "Real exchange rates, imperfect substitutability, and imperfect competition," *Journal of Macroeconomics*, 2007, *29* (4), 639–664.

Maddala, G. S. and Shaowen Wu, "A comparative study of unit root tests with panel data and a new simple test," *Oxford Bulletin of Economics and Statistics*, 1999, *61* (S1), 631–652.

Ng, Serena and Pierre Perron, "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica*, 2001, *69* (6), 1519–1554.

Pedroni, Peter, "Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis," *Econometric Theory*, 2004, *20* (03), 597–625.

Rogoff, Kenneth, "The Purchasing Power Parity Puzzle," *Journal of Economic Literature*, 1996, *19*, 647–668.

Rose, Andrew K., Saktiandi Supaat, and Jacob Braude, "Fertility and the real exchange rate," *Canadian Journal of Economics/Revue canadienne d'économique*, 2009, *42* (2), 496–518.

Samuelson, Paul, "Theoretical Notes on Trade Problems," *Review of Economics and Statistics*, May 1964, *46* (2), 145–154.

Sax, Christoph and Rolf Weder, "How to Explain the High Prices in Switzerland?," *Swiss Journal of Economics and Statistics*, 2009, *145* (IV), 463–483.

Stein, Jerome L. and Polly Reynolds Allen, *Fundamental determinants of exchange rates*, Oxford University Press, USA, 1997.

Stock, James H. and Mark W. Watson, "A simple estimator of cointegrating vectors in higher order integrated systems," *Econometrica*, 1993, *61* (4), 783–820.

Taylor, Mark P., "Purchasing power parity," *Review of International Economics*, 2003, *11* (3), 436–452.

Tica, Josip and Ivo Družić, "The Harrod-Balassa-Samuelson Effect: A Survey of Empirical Evidence," *University of Zagreb, Faculty of Economics and Business, FEB–Working Paper Series (No. 06-7/686)*, 2006.

Chapter 4

Skill-Biased Technological Change and the Real Exchange Rate

Abstract

We sketch a model that shows how skill-biased technological change may reverse the classic Balassa-Samuelson effect, leading to a negative relationship between the productivity in the tradable sector and the real exchange rate. In a small open economy, export goods are produced with capital, high-skilled and low-skilled labor, and traded for imported consumption goods. Non-tradable services are produced with low-skilled labor only. A rise in the productivity of capital has two effects: (1) It may reduce the demand for labor in the tradable sector if the substitutability of low-skilled labor and capital in the tradable sector is high; and (2) it increases the demand for non-tradables and its labor input. Overall demand for low-skilled labor declines if the labor force of the tradable sector is large relative to the labor force of the non-tradable sector. This leads to lower wages and thus to lower prices and a real exchange rate depreciation.

Keywords: Real Exchange Rate, Balassa-Samuelson Hypothesis, Skill-biased Technological Change, General Equilibrium

1 Introduction

The Balassa-Samuelson (BS) hypothesis states that price level differences between countries, expressed in the same currency, can be ascribed to different productivity differentials between the non-tradable and tradable sector. Through wage adjustments in the non-tradable sector, an increase in the productivity of tradables leads to an appreciation of the real exchange rate, while an increase in the productivity of non-tradables has the opposite effect. The hypothesis was simultaneously developed by Balassa (1964) and Samuelson (1964), but has a research precedent in the work of Harrod (1933). It is one of the most widespread explanations for structural deviations from purchasing power parity (Dornbusch, 1985).

There are a number of studies that find evidence supporting the BS hypothesis (see, e.g., De Gregorio and Wolf, 1994; Chinn and Johnston, 1996 or MacDonald and Ricci, 2007) by using panel data on sectoral total factor productivity (TFP). However, all of these studies rely on the discontinued OECD International Sectoral Database (ISDB). When performing a similar analysis with contemporary data, taken from the newly released OECD Productivity Database (PDBi), Gubler and Sax (2011) cannot confirm the hypothesis.

For the last two decades, they find a robust negative relationship between the productivity in the tradable sector and the real exchange rate in the long run, in contrast to BS. Earlier results supporting the BS hypothesis seem to depend strongly on the choice of the data set.[1] The findings of Gubler and Sax (2011) are confirmed once the TFP values are substituted by labor productivity (LP) values from the OECD Structural Analysis (STAN) database. Figure (1) illustrates the negative relationship. The left panel contains the productivity of tradables in relation to the real exchange rate adjusted by country-specific and time-specific effects. For the right panel, the real exchange rate is additionally adjusted by the productivity of non-tradables and the terms of trade. Both estimations with LP data from the STAN database and TFP data from the PDBi show a significant negative relationship.[2]

The fact that there is a robust negative relationship between tradable productivity and the real exchange rate is puzzling. According to the BS hypothesis, a higher productivity in the tradable sector is expected to be associated with a stronger real exchange rate. What causes this puzzle?

This chapter presents a static general-equilibrium model with skill-biased technological change (SBTC). Inspired by the work of De Gregorio and Wolf (1994) and Autor and Dorn (2009), it provides an explanation for the negative relationship between the productivity

[1] The analysis indicates that the discrepancy in the results cannot be ascribed to the change in the sample period.

[2] A detailed analysis reveals that this reversal is mainly driven by the manufacturing sector, i.e., the higher the productivity in manufacturing, the lower is a country's relative price level.

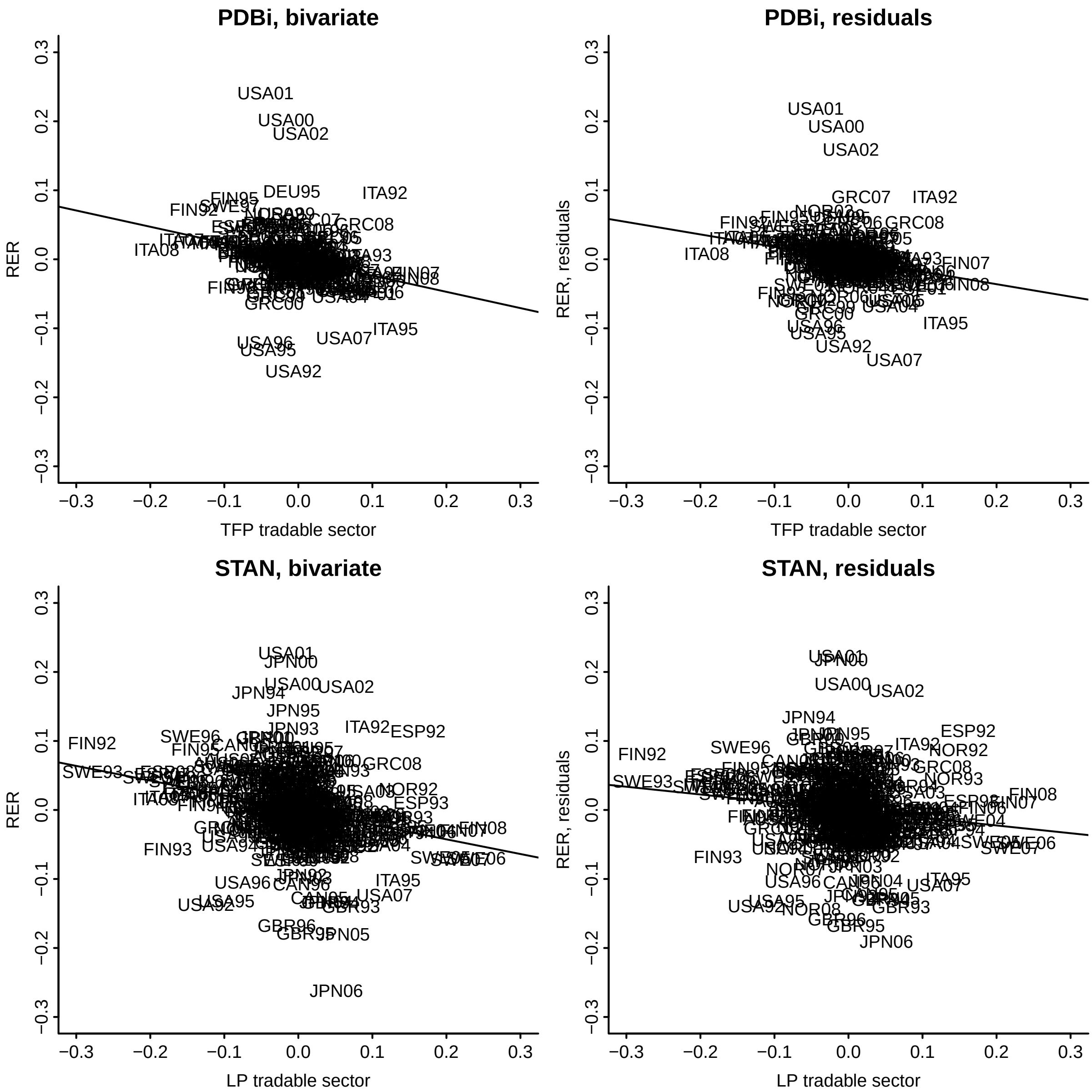


Figure 1: Tradable Productivity and the Real Exchange Rate since 1992

in the tradable sector and the real exchange rate.

Our model shares its basic structure with the model of De Gregorio and Wolf (1994): There is a tradable goods industry that trades its single output good for a single imported good, which is consumed together with a domestically produced non-tradable service.

Furthermore, our model introduces two types of labor, along the lines suggested by Autor and Dorn (2009): low-skilled and high-skilled workers. High-skilled labor is used exclusively in the tradable sector, while low-skilled labor moves freely between the tradable and the non-tradable sector. In the non-tradable sector, low-skilled labor is the only factor of production.

In the tradable sector, low-skilled labor, together with capital, is used to produce an intermediate routine task good, which in turn is combined with high-skilled labor to produce the final tradable good. A key feature of the model is the substitutability of the two factors involved in the production of the intermediate routine task good, low-skilled labor and capital.

In order to analyze the reversion of the BS effect, our study assesses the effect of capital augmenting, i.e. Solow-neutral, technological change on the economy, and especially on the real exchange rate. Ongoing technological progress during the last two decades, particularly in information technology since the 1990s, makes this assumption plausible. Furthermore, Boskin and Lau (2000) identify capital augmenting technological change as the main driver of postwar economic growth of the G7 countries. Alternatively, a very similar effect occurs when the price of capital decreases.[3]

An increase in capital productivity has two effects on the real exchange rate, both operating through their impact on the demand for low-skilled labor. First, under certain conditions, a capital productivity improvement reduces the demand for low-skilled labor in the tradable sector. This is the *labor-repellent effect*. The demand diminishes as long as the elasticity of substitution between low-skilled labor and capital is high relative to the importance of the intermediate routine task good in the production of the final tradable good. We provide the necessary and sufficient condition for the effect to occur.

Second, a rise in capital productivity increases the demand for low-skilled workers in the non-tradable service sector. This is the *labor-attracting effect*. As an increase in capital productivity leads to higher income, consumers can increase their consumption of tradable imported goods. Limited consumer desire to substitute between tradable goods and non-tradable services also increases the demand for non-tradable services, which in turn raises the demand of firms in the non-tradable sector for low-skilled workers.

Depending on whether the labor-repellent effect or the labor-attracting effect is stronger, it is possible that overall demand for low-skilled workers diminishes, i.e., the rise in demand for low-skilled workers in the non-tradable sector does not offset the fall in the demand for these workers to produce tradable goods. Consequently, the wage for low-skilled labor drops in the general equilibrium, because it is assumed that the overall labor force is fixed and the labor market clears. Finally, the price level of the economy decreases. Thus, an increase in tradable productivity may be connected to a lower price level, and leads to an opposite BS effect.

Whether the labor-repellent effect or the labor-attracting effect dominates depends crucially on the fraction of low-skilled labor used in the production of tradable goods. In order to ensure that the labor-repellent effect outweighs the labor-attracting effect for a given wage rate, the labor force in the tradable sector must be large compared to the labor force in the non-tradable sector.

[3] Autor and Dorn (2009) assume a steady fall in the price of capital in their analysis of wage dispersion in the United States.

The remainder of this chapter is organized as follows. We present the structure of the model in Section 2. In Section 3, we discuss the demand for low-skilled labor in the tradable sector. Section 4 derives the demand for low-skilled workers in the non-tradable sector. Overall demand for low-skilled labor is described in Section 5. Section 6 outlines the general equilibrium. Section 7 concludes.

2 Structure of the Economy

The basic structure of the economy is build along the lines suggested by De Gregorio and Wolf (1994): In a small open economy, there are two sectors, each producing a homogeneous good, the tradable exported good, Y_x, and the non-tradable service, Y_n. The tradable good is entirely traded for the imported good, Y_m, at a given world price, p_x. Households gain utility from the consumption of the imported good, Y_m, and the non-tradable service, Y_n. Capital is specific to the tradable sector and assumed to be completely mobile between countries. Low-skilled workers can move between sectors but not between countries. In the following sections, we specify the model in detail.

2.1 Production of Tradables and Non-Tradables

In our model, the production of tradables differs in two ways from the model proposed by De Gregorio and Wolf (1994): First, there is a second type of labor, high-skilled labor, L_h, that is specific to the tradable industry. Second, low-skilled labor, L_x, and capital, K, are close substitutes. Both differences are reflected in a modified production function for tradables that is borrowed from Autor and Dorn (2009):

$$Y_x = L_h^{1-\beta} \left(\underbrace{[(a_r\, L_x)^{\mu} + (a_k\, K)^{\mu}]^{\frac{1}{\mu}}}_{\text{routine task good}} \right)^{\beta} . \tag{1}$$

This Cobb-Douglas production function with $0 < \beta < 1$ nests a Constant Elasticity of Substitution (CES) function, which produces an intermediate routine task good. As capital and low-skilled labor are, by assumption, close substitutes, the elasticity of substitution, $\epsilon = 1/(1-\mu)$, is larger than 1, implying $0 < \mu < 1$. The intermediate routine task good is combined with high-skilled labor to produce the final exported good. $a_r > 0$ and $a_k > 0$ represent exogenous productivity parameters for low-skilled labor and capital, respectively. Note that the productivity parameter for high-skilled labor is normalized to unity, and so a_r and a_k may be interpretated as relative productivity terms.

The production of non-tradables, Y_n, is described by a linear production function in low-skilled labor, L_n, (De Gregorio and Wolf, 1994; Autor and Dorn, 2009):

$$Y_n = a_n \, L_n, \tag{2}$$

where $a_n > 0$ denotes exogenous low-skilled labor productivity in the non-tradable sector.

2.2 Capital and Labor Markets

In our model, we assume that capital is completely mobile between countries. Moreover, the economy is too small to affect the world price of capital. Therefore, firms in the tradable sector can adjust their capital input at a given price $r > 0$.

High-skilled labor is used exclusively in the tradable sector, while low-skilled workers are mobile between the tradable and the non-tradable sector. In the non-tradable sector, low-skilled labor is the only factor of production. We assume that the supply of both low-skilled labor, $\bar{L}_l$, and high-skilled labor, $\bar{L}_h$, is fixed and no transformation from $\bar{L}_l$ to $\bar{L}_h$ is possible. Furthermore, labor cannot move between countries.

2.3 Consumption

Households gain utility from the consumption of the imported good, Y_m, and the non-tradable service, Y_n, according to a CES utility function (De Gregorio and Wolf, 1994; Autor and Dorn, 2009):

$$U = \left({Y_n}^{\phi} + {Y_m}^{\phi} \right)^{\frac{1}{\phi}}, \tag{3}$$

where $\epsilon_c = 1/(1-\phi)$ is the elasticity of substitution between the two consumption goods. We assume that imported goods and non-tradable services are complements, and therefore, $\epsilon_c < 1$, implying $\phi < 0$.

2.4 Prices, Wages, and the Real Exchange Rate

As there is only one international currency, the real exchange rate (RER) between two countries is defined as the ratio of the consumer price index (CPI) of the home country, i, to the CPI of the foreign country, j:

$$\mathrm{RER}_{ij} = \frac{\mathrm{CPI}_i}{\mathrm{CPI}_j}, \tag{4}$$

where the CPI is a weighted average of the goods and services that are consumed domestically, i.e., imported goods and non-tradable services:

$$\mathrm{CPI} = \gamma \, p_n + (1-\gamma) \, p_m, \tag{5}$$

where $\gamma = 1/(1+s_c)$ denotes the share of C_n in total consumption and $s_c = C_m/C_n > 0$ is the fraction of imported goods to non-tradable services.

Without loss of generality, we set the price of the imported good equal to one ($p_m = 1$). Therefore, all prices are expressed in units of the imported good. The normalization has two advantages: First, the price of the exported good, p_x, directly reflects the terms of trade. Second, as the consumer price index is expressed in units of the imported good, Equation (5) simplifies to $\text{CPI} = \gamma\, p_n + (1-\gamma)$. Thus, the price of the non-tradable service, p_n, determines the CPI and the real exchange rate. Finally, p_n is determined by the profit maximizing conditions in the non-tradable service industry: $w = a_n\, p_n$, where w denotes the wage rate and a_n denotes exogenous labor productivity in the non-tradable sector.[4]

In the following, we skip the steps from w to p_n to the real exchange rate, focusing on the behavior of w in the general equilibrium. Once the equilibrium wage rate, w^*, is known, the determination of the equilibrium price of the non-tradable service, p_n^*, the equilibrium consumer price index, CPI*, and the equilibrium real exchange rate, RER*, is straightforward:

$$w^* \Longrightarrow p_n^* \Longrightarrow \text{CPI}^* \Longrightarrow \text{RER}^*$$

3 Low-Skilled Labor Demand in the Tradable Sector

As the supply of low-skilled labor is fixed, the wage rate, w, is determined by the demand for low-skilled workers. Overall demand for low-skilled labor consists of two components: the demand for low-skilled workers in the tradable sector and the demand for low-skilled workers in the non-tradable sector. An exogenous shock, like an increase in capital productivity, may affect the demand for this input factor both in the tradable export sector and in the non-tradable service sector. We will discuss low-skilled labor demand in the tradable sector first.

Given the production function in Equation (1), the profit function is:

$$\pi = p_x\, Y_x - w\, L_x - r\, K - w_h\, L_h, \tag{6}$$

where w and w_h are the real wages of low-skilled labor and high-skilled labor, respectively. Hereafter, low-skilled labor is generally referred to as labor. r denotes the given world real interest rate for capital. Like in any production function with constant returns to scale, the optimal capital intensity in a CES production function does not depend on the level of production. As shown in Appendix 4.A.1, the optimal capital intensity, s, depends only on relative productivities, a_k/a_r, relative factor prices, w/r, and on the elasticity of

[4] The profit maximizing conditions in the non-tradable sector are derived in Section 4.

substitution, μ:

$$s = \frac{K^*}{L_x^*} = \left(\frac{{a_k}^{\mu}\, w}{{a_r}^{\mu}\, r}\right)^{\frac{1}{1-\mu}}. \tag{7}$$

Intuitively, an increase in a_k or a decrease in r makes capital more attractive, causing firms to substitute capital for labor. On the other hand, an increase in a_r or a decrease in w makes labor more attractive, causing firms to substitute labor for capital.

Because s does not depend on the level of L_x and K, we substitute sL_x for K in Equation (1) and replace Y_x in Equation (6) to obtain:

$$\pi = p_x\, L_h^{1-\beta}({a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu})^{\frac{\beta}{\mu}}\, {L_x}^{\beta} - w\, L_x - r\, s\, L_x - w_h\, L_h. \tag{8}$$

As the supply of high-skilled labor is fixed to $\bar{L}_h$, firms will employ all high-skilled workers and optimize over the number of low-skilled workers.[5]

Thus, the first order condition with respect to L_x is:

$$\beta\, p_x\, \bar{L}_h^{1-\beta}({a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu})^{\frac{\beta}{\mu}}\, {L_x}^{\beta-1} = w + r\, s, \tag{9}$$

and has a straightforward interpretation: The left-hand side is the marginal revenue of L_x, taking into account that an increase in L_x also implies a higher K. The right-hand side represents the marginal costs of L_x and the additional amount of K associated with it.

Solving for L_x reveals the optimal demand for labor in the tradable sector:

$$L_x = \bar{L}_h \left(\frac{\beta\, p_x\, ({a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu})^{\frac{\beta}{\mu}}}{w + r\, s}\right)^{\frac{1}{1-\beta}}. \tag{10}$$

Note that capital intensity, s, is itself a function of the parameters a_k, a_r, r, w and μ.

Figure (2) displays the relationship between the wage and low-skilled labor demand in the tradable sector for two values of a_k. As the proof in Appendix 4.A.2 demonstrates, the demand for L_x is decreasing in w for all values in the specified parameter space. Intuitively, there are two reasons behind the decreasing relationship. First, there is a *substitution effect*: An increase in w leads to an increase in s, as firms substitute capital for labor. Second, there is an *income effect.* Even if there was no substitution effect, firms would reduce the number of workers as the optimal level of overall production decreases.

What is the impact of the productivity of capital, a_k, on L_x? Again, an increase in a_k has two effects: Through the substitution effect, a_k negatively affects L_x. With a_k increasing, s rises as firms substitute capital for labor. However, the income effect works in the opposite direction. With an enhanced capital productivity, the optimal production of the final good increases, causing firms to increase their demand for labor. Overall, the

[5]Given the fixed supply $\bar{L}_h$, the wage of high-skilled workers is determined for every level of Y_x. However, for our analysis, the wage of high-skilled workers is irrelevant.

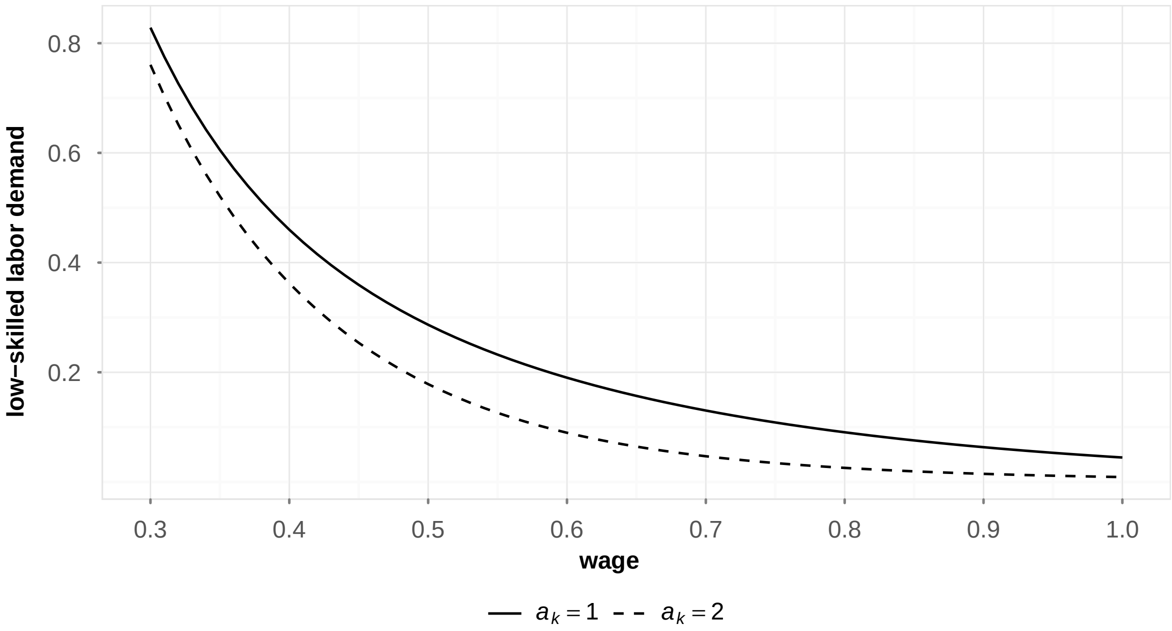


Figure 2: The figure shows a numerical example for low-skilled labor demand in the tradable sector with low capital productivity ($a_k = 1$, solid line), and a numerical example with high capital productivity ($a_k = 2$, dashed line). The other parameters are $a_r = 1$, $\mu = 0.8$, $\beta = 0.5$, $r = 1$, $p_x = 1$, $L_h = 0.3$.

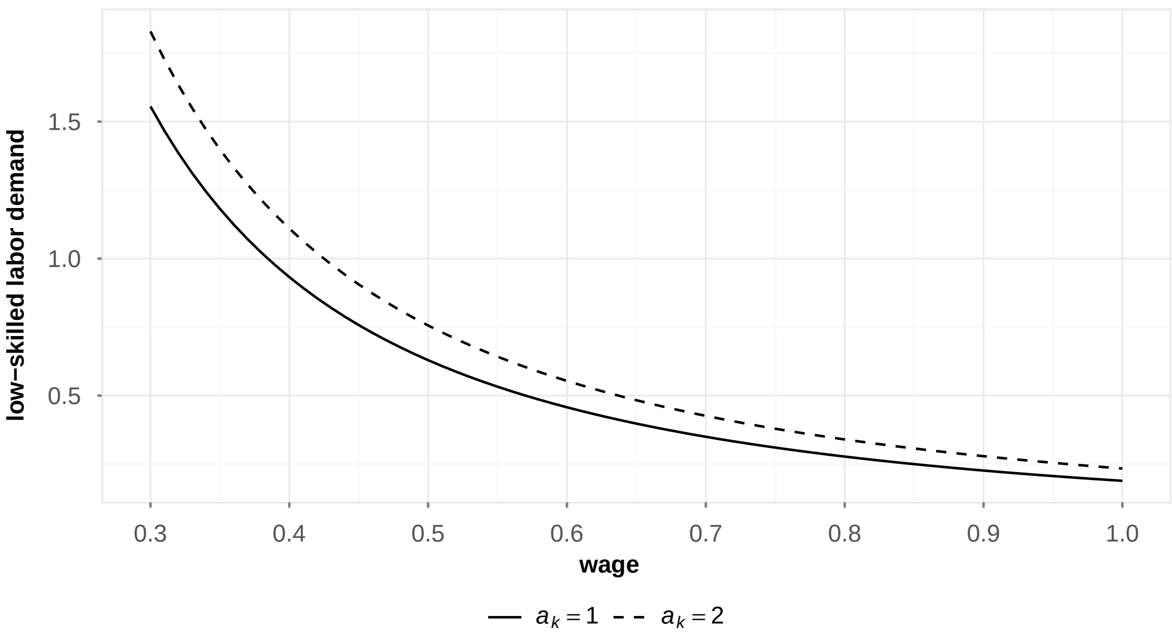


Figure 3: The figure shows a numerical example for low-skilled labor demand in the tradable sector with low capital productivity ($a_k = 1$, solid line), and a numerical example with high capital productivity ($a_k = 2$, dashed line). The other parameters are $a_r = 1$, $\mu = 0.3$, $\beta = 0.5$, $r = 1$, $p_x = 1$, $L_h = 0.3$.

impact of a_k on L_x is ambiguous. As shown in Appendix 4.A.3, the impact of a_k on L_x crucially depends on the relation between μ and β. If and only if $\mu > \beta$, an increase in

a_k leads to a decrease in L_x. In Figure (2), with μ being larger than β, an increase in a_k gives rise to a decrease in L_x, while in Figure (3), with μ being smaller than β, an increase in a_k leads to an increase in L_x.

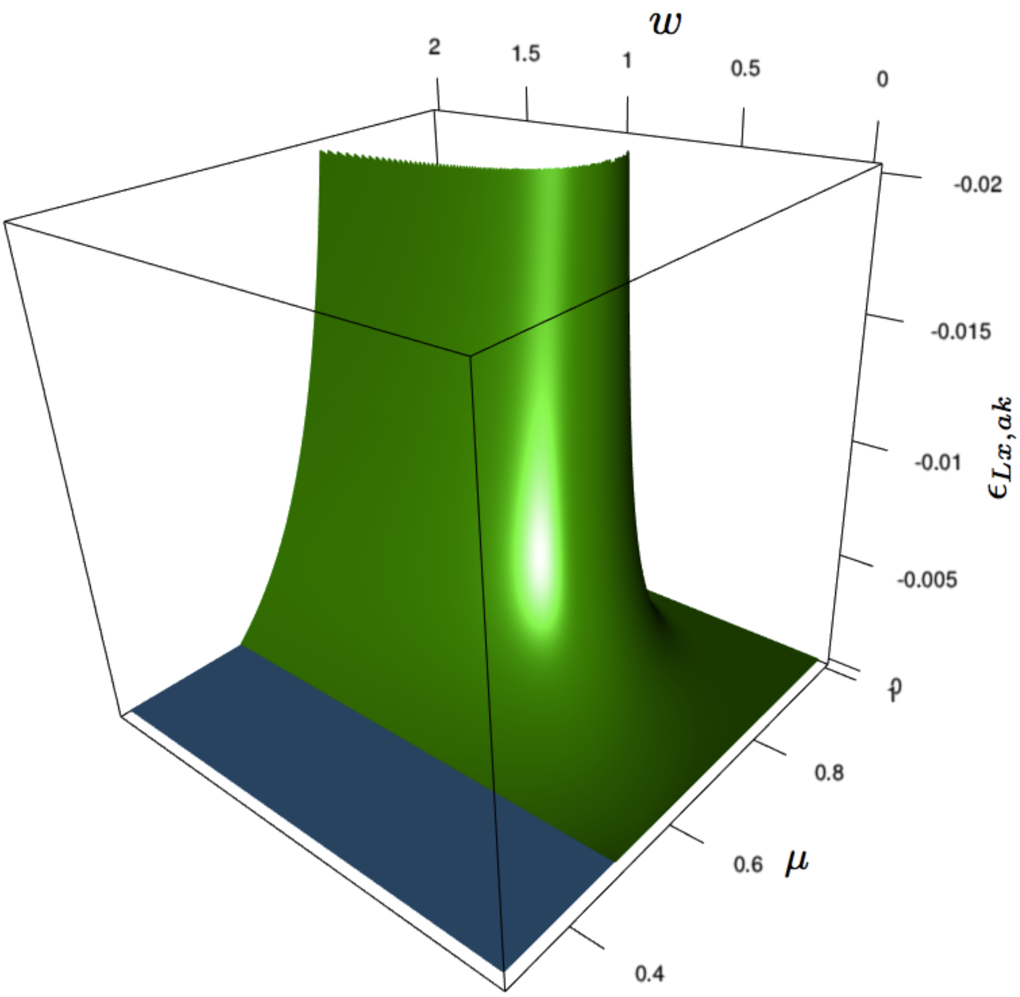


Figure 4: The figure shows the elasticity of low-skilled labor demand in the tradable sector with respect to capital productivity, $\epsilon_{Lx,ak}$, as a function of μ, where $1/(1-\mu)$ is the elasticity of substitution between low-skilled labor and capital, and the wage rate, w. The other parameters are $a_r = 1$, $\beta = 0.5$, $r = 1$, $p_x = 1$.

Intuitively, in order to observe a *labor-repellent effect* in the tradable sector, the substitution effect must be high relative to the income effect. If β is small, the routine task intermediate good has only a limited importance in the production of the final good, and the increase in overall production is small. Thus, β scales the size of the income effect, while μ determines the substitution effect.

The effect of a change in the price of capital, r, on L_x is comparable to the effect of a_k on L_x, but works in the opposite direction. Through the substitution effect, a decrease in

Table 1: The effect of the parameters on L_x

Parameter	Substitution Eff.	Income Effekt	Overall, $\beta < \mu$	Overall, $\beta > \mu$
w	$-$	$-$	$-$	$-$
r	$+$	$-$	$+$	$-$
a_k	$-$	$+$	$-$	$+$
a_r	$+$	$+$	$+$	$+$
L_h			$+$	$+$
p_x			$+$	$+$

r leads to a lower demand for labor; through the income effect, it increases the demand for labor. Overall, if and only if $\mu > \beta$, a decrease in r leads to a decrease in L_x, as shown in Appendix 4.A.4.

Remember that a negative impact of a_k on L_x is a necessary precondition in order to observe the reversed BS effect. Figure (4) depicts the *negative* elasticity of L_x with respect to a_k, $\epsilon_{Lx,ak}$. As soon as the elasticity becomes positive, it is below the surface and not shown. The percentage change in L_x is plotted on the vertical axis, w and μ on the horizontal axes. Figure (4) demonstrates the interplay of μ and w on the elasticity of L_x. From this, the following observations can be made:

First, the elasticity is negative only for values of $\mu > \beta$, as explained above. Second, for values of $\mu > \beta$, there is a transient peak as one moves from low to high values of w: Low wages are associated with low absolute elasticity values, as the number of workers that are substituted by an increase in a_k is very small relative to the size of the labor force in the sector. An increase in w leads to a stronger negative elasticity, but the marginal effect converges to zero, because most of the production of the intermediate routine good is done by capital. Third, if $\mu \to 1$, either labor *or* capital is employed in the production of the intermediate good. At the point where the capital fraction, s, is equal to one, all workers are substituted by capital. Thus, the parameters determining s also determine the shape of the 'elasticity hill'. Note that the figure is truncated at $\epsilon_{Lx,ak} = -0.2$. Therefore, the figure does not show the large negative values for $\mu \to 1$, in order to facilitate the interpretation of the figure.

There are two other parameters that influence L_x. Increasing the supply of high-skilled workers, $\bar{L}_h$, simply scales up the production. The elasticity of L_x with respect to $\bar{L}_h$ is one. As the price of the imported good, p_m is normalized to unity, p_x denotes the terms of trade and has a similar effect as $\bar{L}_h$. A one percent increase in p_x leads to an increase in L_x by $1/(1-\beta)$ percent. Both parameters, $\bar{L}_h$ and p_x, do not affect the elasticity of L_x with respect to a_k, as can be seen from Equation (29) in Appendix 4.A.3. The impact of all parameters on L_x is summarized in Table (1).

The analysis reveals that the relation between the substitution effect determined by

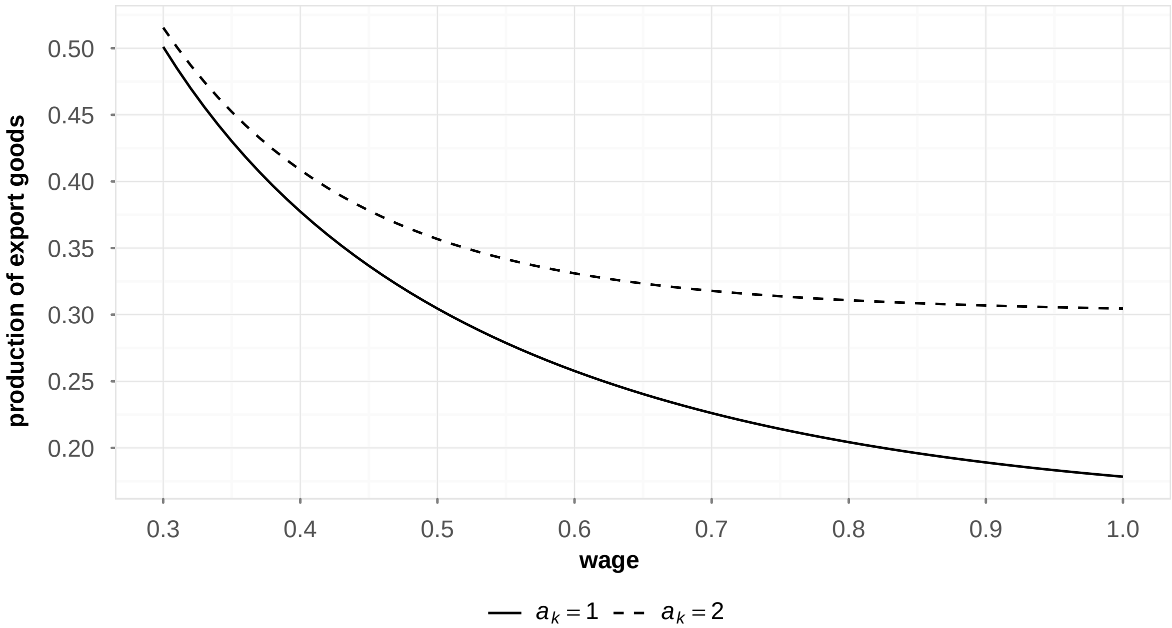


Figure 5: The figure shows a numerical example for the production of exported goods with low capital productivity ($a_k = 1$, solid line), and a numerical example with high capital productivity ($a_k = 2$, dashed line). The other parameters are $a_r = 1$, $\mu = 0.8$, $\beta = 0.5$, $r = 1$, $p_x = 1$, $L_h = 0.3$.

μ and the income effect determined by β is crucial for observing the reverse BS effect. Whether or not the condition is fulfilled is an empirical question. Krusell et al. (2000) estimate the substitution elasticity, $1/(1-\mu)$, between capital and unskilled labor in a four-factor production function to be about 1.7. However, their data set covers both the tradable and non-tradable sector of the U.S. economy, and the sample period ends in 1992. We think that this value is substantially larger for the tradable sector during recent decades, particularly for manufacturing, the largest subsector. Moreover, the income share of capital and low-skilled labor has been falling over the last 30 years, as the relative wage and employment share of high-skilled labor has been rising (Autor and Dorn, 2009).[6]

4 Low-Skilled Labor Demand in the Non-Tradable Sector

The demand for labor in the non-tradable sector is the second component of overall labor demand. An increase in capital productivity affects the demand for labor in the non-tradable sector by increasing the production of exported goods, and by increasing the amount of imported goods available to consumers. Since the elasticity of substitution of the consumers is limited, a higher amount of the imported good leads to an increase in

[6]We assume that the income share of capital remained virtually unchanged.

the demand for non-tradable services, which in turn increases the demand for labor in this sector. We analyze this mechanism step by step.

4.1 Production of Tradable Goods and International Trade

An expression for the production of the exported good, Y_x, can be obtained by inserting the low-skilled labor demand function of the export sector (given in Equation 10) into the production function (given in Equation 1). The resulting function decreases in w, but converges to a constant level associated with capital as the only input in the production of the intermediate routine task good. Figure (5) shows the relationship between the wage and the production of exported goods for two values of a_k. For any given wage, the production of Y_x is increasing in a_k. However, for very low levels of w, almost no capital is employed, and an increase in a_k has only a small effect on production.

In the next step, Y_x is traded for the imported good, Y_m, at the price p_x. Thus, the 'production' of imports by domestic exporters is given by:

$$Y_m = p_x \, Y_x. \tag{11}$$

4.2 Consumers

Because consumers' utility is generated by a CES function (shown in Equation 3), the optimal consumption share, s_c, between the imported good and the non-tradable service is independent of the level of consumption. The same argument that applies to the optimal capital intensity, s, holds for s_c (see Appendix 4.A.1).[7]

Given C_m, p_n and ϕ, the demand for non-tradable services, C_n, is determined by:

$$C_n = s_c \, C_m = \left(\frac{1}{p_n}\right)^{\frac{1}{1-\phi}} C_m. \tag{12}$$

The demand for C_n depends on the relative price of the two consumption goods (which is p_n, as p_m is normalized to unity), the elasticity of substitution, $\epsilon_c = 1/(1-\phi)$, and the demand for tradable goods, C_m.

4.3 Production of the Non-Tradable Service

The non-tradable sector uses labor as its only input factor and linearly transforms it into output. Profit maximizing implies that the wage rate is:

$$w = a_n \, p_n. \tag{13}$$

[7]For the Leontief utility function, a special case of the CES utility function, s_c is equal to one and independent of relative prices.

Substituting p_n in Equation (12) and using the market clearing conditions $Y_m = C_m$ and $Y_n = C_n$, we get:

$$Y_n = \left(\frac{a_n}{w}\right)^{\frac{1}{1-\phi}} Y_m. \tag{14}$$

When the demand for non-tradable services is determined, so is the demand for labor in the non-tradable sector. We make use of Equation (2) to obtain:

$$L_n = a_n{}^{\frac{\phi}{1-\phi}} \left(\frac{1}{w}\right)^{\frac{1}{1-\phi}} Y_m. \tag{15}$$

Because Y_x negatively depends on w, and because Y_m is the product of Y_x and p_x, an increase in w leads to lower imports and to a lower demand for L_n.

How is L_n affected by a_k? For any given wage, an increase in a_k increases the demand for non-tradable labor by raising the amount of imports due to the *labor-attracting effect* $(\partial L_n / \partial a_k > 0)$.

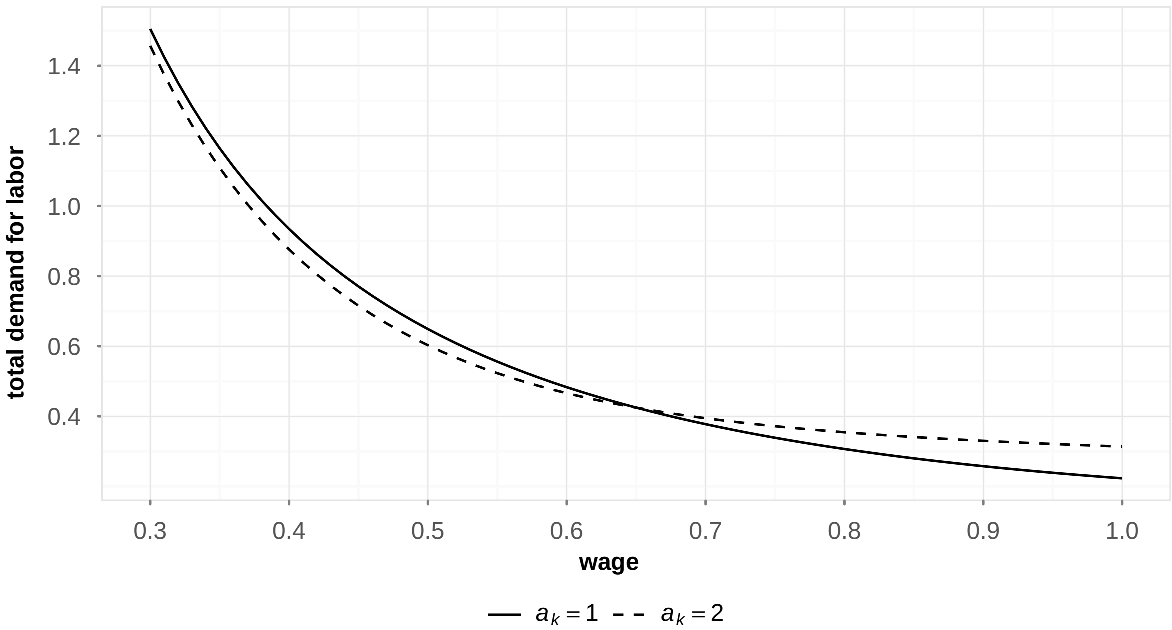


Figure 6: The figure shows a numerical example for low-skilled labor demand in the tradable and the non-tradable sector with a low capital productivity ($a_k = 1$, solid line), and a numerical example with high capital productivity ($a_k = 2$, dashed line). The other parameters are $\phi = -3$, $a_r = 1$, $a_n = 1$, $\mu = 0.8$, $\beta = 0.5$, $r = 1$, $p_x = 1$ and $L_h = 0.3$.

5 Total Demand for Low-Skilled Labor

As a final step, total demand for low-skilled labor is the sum of the demand for labor in both sectors:

$$L_l = L_x + L_n. \tag{16}$$

Figure (6) illustrates the relationship between the wage and total low-skilled labor demand for two values of a_k. The demand for L_l is decreasing in w for all values in the specified parameter space. This is not surprising, as both L_x and L_n are decreasing functions in w. An overview of the effects of the parameters on L_l is given in column (5) of Table (2).

Table 2: The effect of the parameters on s, L_x, Y_x, L_n, L_l and L_n/L_x, given that $\beta < \mu$.

Parameter	s	L_x	Y_x	L_n	L_l	L_n/L_x
w	+	−	−	−	−	+
r	−	+	−	−	−/+ *	−
a_k	+	−	+	+	−/+ *	+
a_r	−	+	+	+	+	−
a_n	∅	∅	∅	−	−	−
ϕ	∅	∅	∅	−	−	−
L_h	∅	+	+	+	+	∅
p_x	∅	+	+	+	+	∅

Notes: * depending on L_n/L_x; ∅: no effect

The impact of a_k on L_l, however, is ambiguous. As it has been shown, $\mu > \beta$ is a necessary precondition in order to observe the labor-repellent effect in the tradable sector ($\partial L_x/\partial a_k < 0$). On the other hand, in the non-tradable sector, an increase in a_k leads to the labor-attracting effect ($\partial L_n/\partial a_k > 0$). If L_n is small compared to L_x, a_k has a negative impact on L_l. In Figure (6), the labor-repellent effect dominates for $w \lessapprox 0.8$, where an increase in a_k has a negative effect on L_l.

Thus, whether the marginal effect of a_k on L_l is positive or negative depends on the relative size of the labor force of the two sectors, L_n/L_x. Column (6) of Table (2) summarizes the effects of the parameters on L_n/L_x. All parameters have an unambiguous effect on L_n/L_x. L_n/L_x negatively depends on r and positively depends on a_k. This follows directly from columns (2) and (4). Because the effects of w and a_r on L_n/L_x are not obvious, proofs are given in Appendices 4.A.5 and 4.A.6. An increase in w leads to an increase in L_n/L_x, while a_r decreases L_n/L_x.

Two other parameters affect the relative labor force of the two sectors: An increase in non-tradable labor productivity, a_n, implies a lower L_n/L_x. Therefore, a_n is positively related with the probability that the labor-repellent effect dominates the labor-attracting effect. If a_n is very large, $L_n/L_x \to 0$, and the L_l-function converges to the L_x-function, as the labor-attracting effect of the non-tradable sector becomes irrelevant relative to the labor-repellent effect of the tradable sector.

The parameter controlling the elasticity of substitution in consumption, ϕ, has a negative but small impact on the relative labor force of the non-tradable sector, L_n/L_x. For reasonable values ($\phi < -2$), $\phi/(1-\phi)$ is already larger than 2/3. As ϕ decreases, the value converges towards one for a Leontief utility function (with $\phi \to -\infty$). In the numerical examples, a value of -3 has been chosen.

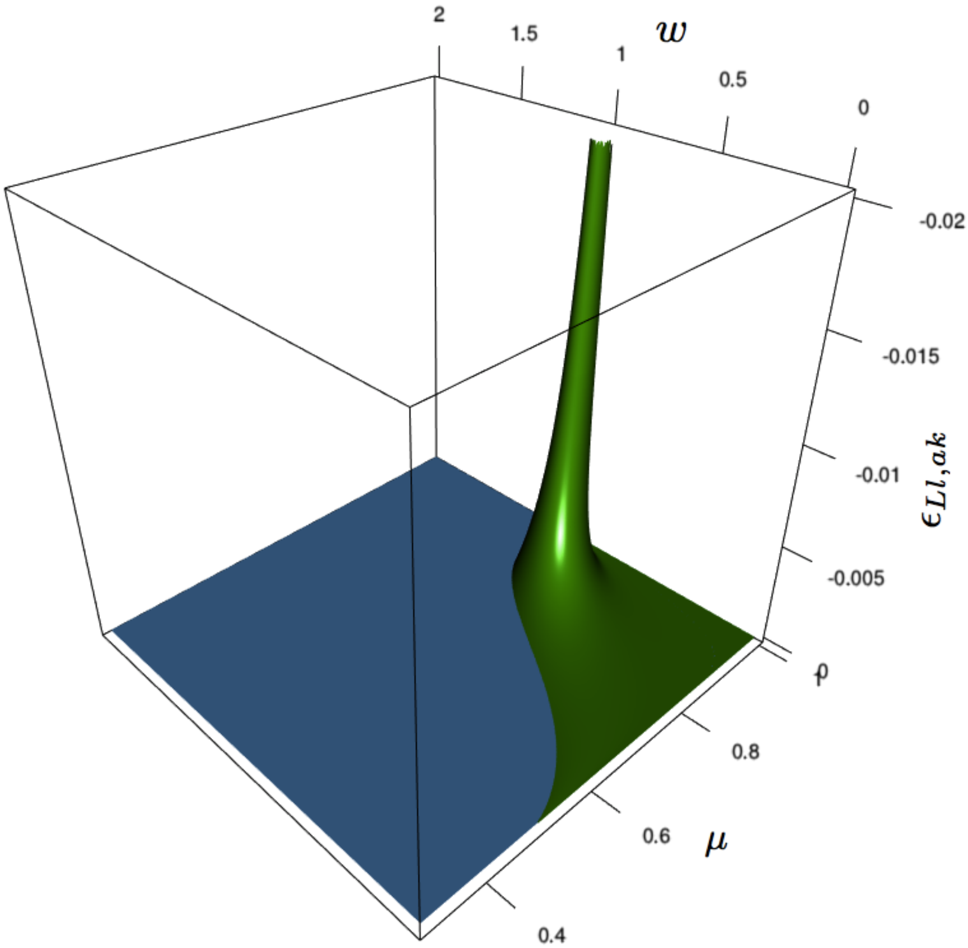


Figure 7: The figure shows the elasticity of low-skilled labor demand in the tradable sector with respect to capital productivity, $\epsilon_{Ll,ak}$, as a function of μ, where $1/(1-\mu)$ is the elasticity of substitution between low-skilled labor and capital, and the wage rate, w. The other parameters are $a_r = 1$, $\beta = 0.5$, $r = 1$, $p_x = 1$.

Figure (7) shows a numerical example that illustrates the reversed BS effect. The figure depicts *negative* values of the elasticity of L_l with respect to a_k, $\epsilon_{Ll,ak}$. As in Figure (4), positive elasticities are below the surface. Again, the percentage change in L_l is plotted on the vertical axis, w and μ on the horizontal axes. We make the following observations:

First, because $\mu > \beta$ is a necessary precondition, negative elasticities can only be observed for $\mu > \beta$. Second, for $\mu > \beta$ the elasticity is negative if $w \to 0$. Intuitively, at low values of w, the intermediate good is produced almost exclusively by labor; an increase in a_k thus has almost no income effect, but a strong substitution effect. As w increases, firms will substitute capital for labor. For some values of w, this leads to a transient increase in the absolute value of the elasticity, as the marginal effect becomes larger relative to the remaining labor force. On the other hand, L_n/L_x increases in w. At some point, the labor-attracting effect of the non-tradable sector dominates the labor-repellent effect, and the elasticity becomes positive, leading to the standard BS result. Third, if $\mu \to 1$, all labor will be substituted at the point where $s = 1$. Comparable to the analysis of L_x in Figure (4), the parameters determining s also determine the shape of the 'elasticity hill'.

As in the case of L_x, there is a one-to-one relationship between $\bar{L}_h$ and L_n. Therefore, the elasticity of L_l with respect to $\bar{L}_h$ is also equal to one. While an increase in $\bar{L}_h$ proportionally increases the marginal effect of a_k on L_l, the elasticity of L_l with respect to a_k is not affected by $\bar{L}_h$. Therefore, $\bar{L}_h$ has no impact on the elasticity function shown in Figure (7). Similarly, p_x does not affect the the elasticity function of L_l with respect to a_k. As in the case of L_x, the elasticity of L_n with respect to p_x is $1/(1-\beta)$. Thus, an increase in p_x does not change the relative size of the sectors.

6 General Equilibrium and the Real Exchange Rate

Recall from Section 2.2 that the supply of labor is fixed and equal to $\bar{L}_l$. Therefore, in equilibrium, the wage rate is determined by setting supply equal to demand:

$$\bar{L}_l = L_l(w^*) = L_x(w^*) + L_n(w^*), \tag{17}$$

where w^* denotes the equilibrium wage rate for low skilled labor. As L_l is decreasing in w, there is a unique solution for w^*. This leads to a positive and monotone relationship between L_l and w^*.

As discussed in Section 2.4, there is a direct link from w^* to the equilibrium price of the non-tradable service, p_n^*, the equilibrium consumer price index, CPI^*, and the equilibrium real exchange rate, RER*.

If an improvement in capital productivity, a_k, diminishes the overall labor demand, L_l, it also decreases w^*, p_n^*, the CPI* and RER*. Therefore, a fall in L_l is sufficient to observe the opposite BS effect. As stated in the previous section, an improvement in a_k decreases L_l, if (1) the substitution effect dominates the income effect in the production of tradable goods, and (2) the labor force of the tradable sector is large relative to the labor force of the non-tradable sector.

7 Summary and Conclusions

We sketch a model that shows how skill-biased technological change may reverse the classic BS effect, leading to a negative relationship between the productivity in the tradable sector and the real exchange rate. In order to find such a relationship, the demand for low-skilled labor in the whole economy must fall in response to a rise in capital productivity. With a fixed supply of labor, this lowers the wage rate of low-skilled workers, and hence, the overall price level and the real exchange rate.

An increase in the productivity of capital has two effects on low-skilled labor demand: (1) a *labor-attracting* effect in the non-tradable sector and (2) a (potential) *labor-repellent* effect in the tradable sector. First, an increase in productivity leads to a higher income level in the whole economy. As consumers spend additional income in both the non-tradable and the tradable sector, the demand for low-skilled workers in the non-tradable sector increases. Second, an increase in the productivity of capital potentially decreases the demand for labor in the tradable sector. Such a negative effect occurs if and only if the substitution elasticity between low-skilled labor and capital is high relative to the importance of the routine task good in the production of the final good.

In order to observe the opposite BS effect, the labor-repellent effect in the tradable sector must outweigh the labor-attracting effect in the non-tradable sector. For the labor-repellent effect to dominate the labor-attracting effect, the low-skilled labor force in the tradable sector must be large compared to the non-tradable sector labor force. If the labor force in the tradable sector is small relative to the non-tradable sector, the labor-repellent effect is dominated by the labor-attracting effect, and the classic BS effect occurs.

Several testable hypotheses can be derived from our model: According to our model, the opposite BS effect should be observed in countries where (1) capital productivity enhancement dominates low-skilled labor productivity gains, (2) the income share of high-skilled labor is high (low β), (3) capital is substitutable for low-skilled labor, and (4) the labor force of the tradable sector is large relative to the labor force of the non-tradable sector.

In the United States, for example, employment in low-skilled occupations in industry and agriculture has strongly decreased over the last decades (Autor and Dorn, 2009). Gains in productivity, thus, have led to a decrease in demand for low-skilled workers in the tradable industry. At the same time, an increase in the demand for non-tradable services has led to a strong increase in non-tradable service occupations between 1980 and 2005 (Autor and Dorn, 2009). While the wage growth in these service occupations was stronger than the wage growth in similarly low-skilled occupations in industry, the overall effect on relative wages for low-skilled workers was clearly negative. According to our model, this has reduced the relative costs of non-tradable goods, leading to a depreciation of the real exchange rate, in line with the empirical findings of Gubler and Sax (2011). As we have stated, the occurrence of the reversed BS effect is temporary. Today, productivity gains in US-industry may well lead to the traditional BS effect and a real exchange apreciation. This is because the share of low-skilled labor in industry has become small. Productivity enhancements thus are likely to increase tradable production (a large income effect) while only a small number of workers are laid off (a small substitution effect).

We do not expect the revese BS effect to be of major importance in emerging economies such as China. This is because low-skilled labor is still the dominant factor in tradable production. At current Chinese wage rates, capital intensity in the production of routine tasks is low. An increase in capital productivity would have only a very small negative effect on the low-skilled labor force in the tradable sector.

Of course, at this stage, the model provides only one possible explanation for the empirical finding of an opposite BS effect in the tradable sector. Ultimately, the model needs to be tested empirically. Future research should include further exploration of the hypotheses and their validation.

Appendix 4.A Mathematical Appendix

4.A.1 Optimal Capital Intensity s

We start with the production function:

$$Y_x = L_h^{1-\beta}[(a_r\, L_x)^\mu + (a_k\, K)^\mu]^{\frac{\beta}{\mu}}. \tag{18}$$

The first order conditions with respect to L_x and K are:

$$\frac{L_h^{1-\beta}\beta\, p_x\, (a_r\, L_x^*)^\mu\, ((a_r\, L_x^*)^\mu + (a_k\, K^*)^\mu)^{\frac{\beta}{\mu}-1}}{L_x^*} = w \tag{19}$$

and

$$\frac{L_h^{1-\beta}\beta\, p_x\, (a_k\, K^*)^\mu\, ((a_r\, L_x^*)^\mu + (a_k\, K^*)^\mu)^{\frac{\beta}{\mu}-1}}{K^*} = r. \tag{20}$$

Dividing Equation (19) by Equation (20) yields:

$$\frac{(a_r\, L_x^*)^\mu\, K^*}{(a_k\, K^*)^\mu\, L_x^*} = \frac{w}{r}. \tag{21}$$

Define $s = K^*/L_x^*$ and solve for s to obtain Equation (7).

4.A.2 L_x is Decreasing in w

Proposition. *For the parameters a_k, a_r, r, w, $p_x > 0$, $0 < \mu < 1$ and $0 < \beta < 1$: L_x is a decreasing function of w.*

Proof. Taking the logarithm of both sides of Equation (10) and differentiating with respect to w yields:

$$\frac{\partial \log L_x}{\partial\, w} = \frac{1}{1-\beta}\left[\frac{\beta}{\mu}\; \frac{{a_k}^\mu \mu s^{\mu-1}\frac{\partial s}{\partial w}}{{a_k}^\mu\, s^\mu + {a_r}^\mu} - \frac{1 + r\frac{\partial s}{\partial w}}{w + r\, s}\right]. \tag{22}$$

The derivative of s with respect to w is:

$$\frac{\partial s}{\partial\, w} = \frac{\left(\frac{{a_k}^\mu\, w}{{a_r}^\mu\, r}\right)^{\frac{1}{1-\mu}}}{w\,(1-\mu)} = \frac{s}{w\,(1-\mu)}. \tag{23}$$

Substituting this result in Equation (22) yields:

$$\frac{\partial \log L_x}{\partial\, w} = \frac{1}{1-\beta}\left[\frac{\beta\, {a_k}^\mu\, s^\mu}{w\,(1-\mu)\,({a_k}^\mu\, s^\mu + {a_r}^\mu)} - \frac{r\, s + (1-\mu)w}{w(1-\mu)(w + r\, s)}\right]. \tag{24}$$

For $\partial\, L_x/\partial\, w < 0$, the term in the square bracket must be negative, therefore:

$$\frac{\beta\, {a_k}^\mu\, s^\mu}{{a_k}^\mu\, s^\mu + {a_r}^\mu} < \frac{r\, s + (1-\mu)w}{w + r\, s}. \tag{25}$$

We multiply both sides of Equation (25) by $(w + r\, s)$ and $({a_k}^\mu\, s^\mu + {a_r}^\mu)$ to obtain:

$$\beta\, {a_k}^\mu\, s^\mu (w + r\, s) < (r\, s + (1-\mu)w)({a_k}^\mu\, s^\mu + {a_r}^\mu). \tag{26}$$

Subtracting $(\beta\, {a_k}^{\mu}\, s^{\mu}(w + r\, s))$ on both sides and rearranging yields:

$$\begin{aligned} 0 < (1-\mu)w\, {a_r}^{\mu} + r\, s\, {a_r}^{\mu} + (1-\beta)\, {a_k}^{\mu}\, s^{1+\mu}\, r \\ + (1-\beta)\, {a_k}^{\mu}\, s^{\mu}\, w - {a_k}^{\mu}\, s^{\mu}\, w\, \mu. \end{aligned} \tag{27}$$

We replace ${a_k}^{\mu}$ by $(s^{1-\mu}\, {a_r}^{\mu}\, r)/w$ (see Equation 7) in the last term of the right-hand side and rearrange to obtain:

$$\begin{aligned} 0 < (1-\mu)w\, {a_r}^{\mu} + (1-\beta)\, {a_k}^{\mu}\, s^{1+\mu}\, r + (1-\beta)\, {a_k}^{\mu}\, s^{\mu}\, w \\ + (1-\mu) r\, s\, {a_r}^{\mu}. \end{aligned} \tag{28}$$

Since $0 < \beta < 1$ and $0 < \mu < 1$, the right-hand side is positive and $\partial\, L_x / \partial\, w < 0$.

□

4.A.3 Necessary and Sufficient Condition for $\partial\, L_x / \partial\, a_k < 0$

Proposition. *For the parameters a_k, a_r, r, w, $p_x > 0$, $0 < \mu < 1$ and $0 < \beta < 1$: $1 > \mu > \beta$ is a necessary and sufficient condition for $\frac{\partial \log L_x}{\partial\, a_k} < 0$.*

Proof. Taking the logarithm of both sides of Equation (10) and differentiating with respect to a_k, one gets:

$$\frac{\partial \log L_x}{\partial\, a_k} = \frac{1}{1-\beta}\left[\frac{\beta}{\mu}\, \frac{{a_k}^{\mu} \mu s^{\mu-1} \frac{\partial s}{\partial a_k} + \mu {a_k}^{\mu-1} s^{\mu}}{{a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu}} - \frac{r \frac{\partial s}{\partial a_k}}{w + r\, s}\right]. \tag{29}$$

The derivative of s with respect to a_k is:

$$\frac{\partial s}{\partial a_k} = \frac{\mu \left(\frac{{a_k}^{\mu}\, w}{{a_r}^{\mu}\, r}\right)^{\frac{1}{1-\mu}}}{a_k\, (1-\mu)} = \frac{\mu\, s}{a_k\, (1-\mu)}. \tag{30}$$

Substituting this result in Equation (29) yields:

$$\frac{\partial \log L_x}{\partial\, a_k} = \frac{1}{1-\beta}\left[\frac{\beta}{\mu}\, \frac{\frac{{a_k}^{\mu-1} \mu^2 s^{\mu}}{1-\mu} + \mu {a_k}^{\mu-1} s^{\mu}}{{a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu}} - \frac{\mu\, r\, s}{a_k (1-\mu)(w + r\, s)}\right]. \tag{31}$$

$\frac{\partial \log L_x}{\partial\, a_k} < 0$ only holds if the square bracket in Equation (31) is negative:

$$\frac{\beta}{\mu}\, \frac{\frac{{a_k}^{\mu-1} \mu^2 s^{\mu}}{1-\mu} + \mu {a_k}^{\mu-1} s^{\mu}}{{a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu}} < \frac{\mu\, r\, s}{a_k (1-\mu)(w + r\, s)}. \tag{32}$$

After multiplying both sides of Equation (32) by $(1-\mu)({a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu})$, we obtain:

$$\frac{\beta}{\mu}\, ({a_k}^{\mu-1} \mu^2 s^{\mu} + (1-\mu)\mu {a_k}^{\mu-1} s^{\mu}) < \frac{\mu\, r\, s({a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu})}{a_k (w + r\, s)}. \tag{33}$$

After some manipulations and cancelling μ on both sides, we get:

$$\frac{\beta}{\mu}\, {a_k}^{\mu} s^{\mu-1} < \frac{r({a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu})}{(w + r\, s)}. \tag{34}$$

Multiplying both sides of Equation (34) by $(w + r\,s)$ yields:

$$\frac{\beta}{\mu}\; w\,{a_k}^{\mu} s^{\mu-1} + \frac{\beta}{\mu}\; r\,{a_k}^{\mu} s^{\mu} < r\,{a_k}^{\mu} s^{\mu} + r\,{a_r}^{\mu}. \tag{35}$$

We replace ${a_r}^{\mu}$ by $({a_k}^{\mu}\,w)/(s^{1-\mu}\,r)$ (see Equation 7) in the last term of the right-hand side to get:

$$\frac{\beta}{\mu}\; w\,{a_k}^{\mu} s^{\mu-1} + \frac{\beta}{\mu}\; r\,{a_k}^{\mu} s^{\mu} < r\,{a_k}^{\mu} s^{\mu} + w\,{a_k}^{\mu} s^{\mu-1}. \tag{36}$$

We subtract $(r\,{a_k}^{\mu} s^{\mu} + w\,{a_k}^{\mu} s^{\mu-1})$ from both sides and rearrange to obtain:

$$\left[\frac{\beta}{\mu} - 1\right]\;(w\,{a_k}^{\mu} s^{\mu-1} + r\,{a_k}^{\mu} s^{\mu}) < 0. \tag{37}$$

Therefore, if and only if $\mu > \beta$, Equation (37) holds, and thus $\frac{\partial \log L_x}{\partial\, a_k} > 0$. □

4.A.4 Necessary and Sufficient Condition for $\partial\, L_x/\partial\, r > 0$

Proposition. *For the parameters a_k, a_r, r, w, $p_x > 0$, $0 < \mu < 1$ and $0 < \beta < 1$, $1 > \mu > \beta$ is a necessary and sufficient condition for $\frac{\partial \log L_x}{\partial\, r} > 0$.*

Proof. In order to draw on Proof 4.A.2, we show that $1 > \mu > \beta$ is a necessary and sufficient condition for $\frac{\partial \log K}{\partial\, w} > 0$. Without loss of generality we can redefine s as $s = L_x/K$. Then, the same result applies to $\frac{\partial \log L_x}{\partial\, r}$.

We take the logarithm of both sides of $K = sL_x$ and differentiate with respect to w to obtain:

$$\frac{\partial \log K}{\partial\, w} = \frac{\partial \log s}{\partial\, w} + \frac{\partial \log L_x}{\partial\, w}. \tag{38}$$

From Proof 4.A.2 we know the result of $\frac{\partial \log L_x}{\partial\, w}$, given in Equation (24). The derivative of $\log s$ with respect to w yields:

$$\frac{\partial \log s}{\partial\, w} = \frac{1}{(1-\mu)\,w}. \tag{39}$$

Therefore, substituting the results of Equation (24) and Equation (39) in Equation (38) gives:

$$\begin{aligned}\frac{\partial \log K}{\partial\, w} = \frac{1}{(1-\mu)\,w} &+ \frac{1}{1-\beta}\,\frac{\beta\,{a_k}^{\mu}\,s^{\mu}}{w\,(1-\mu)\,({a_k}^{\mu}\,s^{\mu} + {a_r}^{\mu})} \\ &- \frac{1}{1-\beta}\,\frac{r\,s + (1-\mu)w}{w(1-\mu)(w + r\,s)}.\end{aligned} \tag{40}$$

In order to get $\frac{\partial \log K}{\partial\, w} > 0$, it must be that:

$$\frac{1}{(1-\mu)\,w} + \frac{1}{1-\beta}\,\frac{\beta\,{a_k}^{\mu}\,s^{\mu}}{w\,(1-\mu)\,({a_k}^{\mu}\,s^{\mu} + {a_r}^{\mu})} > \frac{1}{1-\beta}\,\frac{r\,s + (1-\mu)w}{w(1-\mu)(w + r\,s)}. \tag{41}$$

After some manipulations we obtain:

$$0 > {a_r}^{\mu}\;\left[(\beta - \mu)\,(w + rs)\right]. \tag{42}$$

Therefore, $1 > \mu > \beta$ is a necessary and sufficient condition for this equation to hold. $\square$

4.A.5 L_n/L_x is Increasing in w

Proposition. *For the parameters a_k, a_r, r, w, $p_x > 0$, $\phi < 0$, $0 < \mu < 1$ and $0 < \beta < 1$: L_n/L_x is an increasing function of w.*

Proof. We show that $\frac{\partial \log L_x}{\partial w} < \frac{\partial \log L_n}{\partial w}$. From Proof 4.A.2 we know $\frac{\partial \log L_x}{\partial w}$. Taking the logarithm of both sides of Equation (14) and differentiating with respect to w yields:

$$\begin{aligned} \frac{\partial \log L_n}{\partial w} &= -\frac{1}{w(1-\phi)} + \frac{\partial \log Y_x}{\partial w}, \\ \frac{\partial \log L_n}{\partial w} &= -\frac{1}{w(1-\phi)} + \beta\,\frac{\partial \log L_x}{\partial w} + \frac{\beta}{\mu}\,\frac{\partial \log\left(a_r^{\mu} + a_k^{\mu} s^{\mu}\right)}{\partial w}. \end{aligned} \tag{43}$$

Therefore, the inequality $\frac{\partial \log L_x}{\partial w} < \frac{\partial \log L_n}{\partial w}$ is equal to:

$$\frac{\partial \log L_x}{\partial w} < -\frac{1}{w(1-\phi)} + \beta\,\frac{\partial \log L_x}{\partial w} + \frac{\beta}{\mu}\,\frac{\partial \log\left(a_r^{\mu} + a_k^{\mu} s^{\mu}\right)}{\partial w}. \tag{44}$$

By subtracting $\beta\,\frac{\partial \log L_x}{\partial w}$ from both sides and using the result from Equation (24), we get:

$$\begin{aligned} \frac{\beta\, {a_k}^{\mu}\, s^{\mu}}{w\,(1-\mu)\,({a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu})} - \frac{r\,s + (1-\mu)w}{w\,(1-\mu)(w + r\,s)} &< -\frac{1}{w(1-\phi)} \\ &\quad + \frac{\beta\, {a_k}^{\mu}\, s^{\mu}}{w\,(1-\mu)\,({a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu})}. \end{aligned} \tag{45}$$

After some manipulations we obtain:

$$0 < \frac{\mu\, r\, s - w\,\phi - r\,s\,\phi - \mu\, w\,\phi}{1-\phi}. \tag{46}$$

This equation holds given the parameter value restrictions. $\square$

4.A.6 L_n/L_x is Decreasing in a_r

Proposition. *For the parameters a_k, a_r, r, w, $p_x > 0$, $\phi < 0$, $0 < \mu < 1$ and $0 < \beta < 1$: L_n/L_x is a decreasing function of a_r.*

Proof. We show that $\frac{\partial \log L_x}{\partial a_r} > \frac{\partial \log L_n}{\partial a_r}$. Taking the logarithm of both sides of Equation (10) and differentiating with respect to a_r yields:

$$\frac{\partial \log L_x}{\partial a_r} = \frac{1}{1-\beta}\left[\frac{\beta}{\mu}\,\frac{\mu {a_r}^{\mu-1} + {a_k}^{\mu}\mu s^{\mu-1}\frac{\partial s}{\partial a_r}}{{a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu}} - \frac{r\,\frac{\partial s}{\partial a_r}}{w + r\,s}\right]. \tag{47}$$

The derivative of s with respect to a_r is:

$$\frac{\partial s}{\partial a_r} = \frac{-\mu\left(\frac{{a_k}^{\mu}\, w}{{a_r}^{\mu}\, r}\right)^{\frac{1}{1-\mu}}}{a_r\,(1-\mu)} = \frac{-\mu\, s}{a_r\,(1-\mu)}. \tag{48}$$

Substituting this result in Equation (47) yields:

$$\frac{\partial \log L_x}{\partial\, a_r} = \frac{1}{1-\beta}\left[\frac{\beta}{\mu}\,\frac{\mu {a_r}^{\mu-1} - \frac{{a_k}^{\mu}\mu^2 s^{\mu}}{a_r(1-\mu)}}{{a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu}} + \frac{r\, s\, \mu}{a_r(1-\mu)(w+r\, s)}\right]. \tag{49}$$

a_r affects Y_n only via the output of the exported good Y_x (see Equation 14). Hence, taking the logarithm of both sides of Equation (1) with K replaced by sL_x and differentiating with respect to a_r yields:

$$\frac{\partial \log Y_x}{\partial\, a_r} = \beta\,\frac{\partial \log L_x}{\partial\, a_r} + \frac{\beta}{\mu}\,\frac{\partial \log\left(a_r^{\mu} + a_k^{\mu} s^{\mu}\right)}{\partial\, a_r}. \tag{50}$$

Therefore, the inequality $\frac{\partial\ \log L_x}{\partial\, a_r} > \frac{\partial \log L_n}{\partial\, a_r}$ is equal to:

$$\frac{\partial \log L_x}{\partial\, a_r} > \beta\,\frac{\partial \log L_x}{\partial\, a_r} + \frac{\beta}{\mu}\,\frac{\partial \log\left(a_r^{\mu} + a_k^{\mu} s^{\mu}\right)}{\partial\, a_r}. \tag{51}$$

By subtracting $\beta\,\frac{\partial \log L_x}{\partial\, a_r}$ from both sides and using the result from Equation (49), we get:

$$\frac{\beta}{\mu}\,\frac{\mu {a_r}^{\mu-1} - \frac{{a_k}^{\mu}\mu^2 s^{\mu}}{a_r(1-\mu)}}{{a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu}} + \frac{r\, s\, \mu}{a_r(1-\mu)(w+r\, s)} > \tag{52}$$

$$\frac{\beta}{\mu}\,\frac{\mu {a_r}^{\mu-1} - \frac{{a_k}^{\mu}\mu^2 s^{\mu}}{a_r(1-\mu)}}{{a_k}^{\mu}\, s^{\mu} + {a_r}^{\mu}}. \tag{53}$$

Therefore, we obtain:

$$\frac{r\, s\, \mu}{a_r(1-\mu)(w+r\, s)} > 0. \tag{54}$$

This equation holds, given the parameter value restrictions. □

Bibliography

Autor, David H. and David Dorn, "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market," NBER Working Paper 15150, National Bureau of Economic Research, Cambridge, MA July 2009.

Balassa, Bela, "The Purchasing-Power Parity Doctrine: A Reappraisal," *The Journal of Political Economy*, December 1964, *72* (6), 584–596.

Boskin, Michael J. and Lawrence J. Lau, "Generalized Solow-Neutral Technical Progress and Postwar Economic Growth," NBER Working Paper 8023, National Bureau of Economic Research, Cambridge, MA December 2000.

Chinn, Menzie D. and Louis Johnston, "Real Exchange Rate Levels, Productivity and Demand Shocks: Evidence from a Panel of 14 Countries," NBER Working Paper 5709, National Bureau of Economic Research, Cambridge, MA August 1996.

De Gregorio, José and Holger C. Wolf, "Terms of Trade, Productivity, and the Real Exchange Rate," NBER Working Paper 4807, National Bureau of Economic Research, Cambridge, MA July 1994.

Dornbusch, Rudiger, "Purchasing Power Parity," NBER Working Paper 1591, National Bureau of Economic Research, Cambridge, MA March 1985.

Gubler, Matthias and Christoph Sax, "The Balassa-Samuelson Effect Reversed: New Evidence from OECD Countries," WWZ Discussion Paper 09, University of Basel December 2011.

Harrod, Roy F., *International Economics* Cambridge Economic Handbooks, London: Nisbet & Cambridge University Press, 1933.

Krusell, Per, Lee E. Ohanian, José-Víctor Ríos-Rull, and Giovanni L. Violante, "Capital-skill complementarity and inequality: A macroeconomic analysis," *Econometrica*, 2000, *68* (5), 1029–1053.

MacDonald, Ronald and Luca Antonio Ricci, "Real exchange rates, imperfect substitutability, and imperfect competition," *Journal of Macroeconomics*, 2007, *29* (4), 639–664.

Samuelson, Paul, "Theoretical Notes on Trade Problems," *Review of Economics and Statistics*, May 1964, *46* (2), 145–154.

Curriculum Vitae

Matthias Gubler was born on December 23rd, 1980 in Breitenbach (Solothurn, Switzerland) as the first child of Hans-Rudolf Gubler and Monica Gubler.

From 1987 to 1992, Matthias attended elementary school in Breitenbach and subsequently secondary school in Laufen, Baselland, until 1995. From 1995 to 1999, Matthias attended the Gymnasium in Laufen, completing his Matura (high school diploma) with a major in natural sciences.

After working as a trainee at UBS AG in Basel and a language stay in Cairns, Australia, during 2000, Matthias began his studies in economics at the Faculty of Business and Economics of the University of Basel the same year. He continued to work part-time for UBS AG and later for Habasit AG in Reinach, Baselland. In January 2003, Matthias interrupted his studies for a trainee program at Credit Suisse in Zürich, Switzerland, and a language stay in Madrid, Spain, before returning to the Faculty of Business and Economics of the University of Basel in fall 2003. In 2004, he received his Bachelor of Arts in Business and Economics. Matthias continued his studies in Basel and graduated in 2006, receiving his Master of Science in Business and Economics with a major in Monetary Economics and Financial Markets.

From mid-2006 until early 2007, Matthias worked for Habasit AG as a staff member of Group Controlling and Customer Relationship Marketing. Afterwards, he successfully completed the "Swiss Program for Beginning Doctoral Students in Economics" at the Study Center Gerzensee, Switzerland, of the Swiss National Bank in 2007. From 2008 to September 2011, he was employed as a research and teaching assistant for Prof. Dr. Peter Kugler in the Department of Monetary Macroeconomics at the Faculty of Business and Economics of the University of Basel. During his assignments at the university, Matthias wrote his Ph.D. thesis, which he successfully completed in April 2012 with distinction (summa cum laude). In addition, Matthias attended selected advanced doctoral courses in Gerzensee, Zürich and Barcelona, Spain. From April 2010 until September 2011, he was also employed part-time for Novartis AG in Basel as an economist.

Since September 2011, Matthias is working as an economist in the Economic Analysis Unit of the Swiss National Bank.